

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: July 5, 2006, 09:47:58 ; Search time 107 Seconds
(without alignments)
3473.990 Million cell updates/sec

Title: US-10-509-950-1

Perfect score: 4235
Sequence: 1 MSEGQDNLQAIABEGTGEQ.....SASCRTNCASMSNMGQVYK 813

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A-Geneseq2.8:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*
10: geneseqp2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4235	100.0	813	7	ADD89836 Human CAM
2	4235	100.0	813	7	ADG17598 Human Brl
3	4208	99.4	847	7	ADG17594 Human TAR
4	4111	97.1	793	7	ADG17600 Human Brl
5	4084	96.4	827	7	ADG17596 Human Brl
6	3586	84.7	682	5	ABP43676 DKFZP566n
7	3421	80.8	807	7	ADG17603 Mouse TAR
8	3414.5	80.6	651	7	ADG17605 Human TAR
9	3398.5	80.2	648	7	ADG17604 Human TAR
10	2513.5	59.4	645	4	ABG04987 Human pro
11	2047	48.3	438	4	ABG02090 Novel hum
12	1737.5	41.0	1044	8	ADRO9546 Human pro
13	1510	35.7	971	6	ADDE56304 Human pro
14	1493	35.3	976	6	ABRS8323 BC00147B
15	1455	34.4	1027	6	ABR43214 Human TRA
16	1451	34.3	958	4	ABRS8322 Human TRA
17	1420.5	33.5	433	4	ABG02088 Novel hum
18	1332.5	31.5	840	7	ADG17199 Nuclear f
19	1294.5	30.6	854	7	ADG17201 Nuclear f
20	1020	24.1	366	4	ABG04986 Novel hum
21	944	22.3	371	4	ABG02089 Novel hum
22	760.5	18.0	723	7	ADG17197 Nuclear f
23	582	13.7	123	4	ABBI5877 Human ner

ALIGNMENTS

24	491	11.6	162	4	ABBI7003	Abbi7003 Human ner
25	473	11.2	107	7	ADB65582	ADB65582 Human pro
26	430	10.2	89	5	AAW47778	AAW47778 Human har
27	430	10.2	89	7	ADG17606	ADG17606 Human pro
28	430	10.2	89	7	ADG17602	ADG17602 Human Brl
29	423	10.0	89	3	ABBI01375	Abbi01375 Neuron-as
30	410.5	9.7	1805	4	ABG65262	Abg65262 Drosophi1
31	346.5	8.2	94	4	ABG04984	Abg04984 Novel hum
32	346.5	8.2	94	4	ABG02085	Abg02085 Novel hum
33	344.5	8.1	88	7	ADG17607	ADG17607 Mouse pro
34	311.5	7.4	78	5	ABP33783	Abp33783 Human ORF
35	308	7.3	79	4	ABG04985	Abg04985 Novel hum
36	308	7.3	79	4	ABG02086	Abg02086 Novel hum
37	293	6.9	2004	5	ABG95113	Abg95113 Human tra
38	293	6.9	2004	9	ADX07185	Adx07185 Cyclin-de
39	293	6.9	2004	9	ADY17185	Ady17185 PRO polyP
40	291.5	6.9	1400	9	ADV85501	Adv85501 Mouse bro
41	291	6.9	2000	9	ADY14779	Ady14779 PRO polyP
42	291	6.9	2703	4	ABB63299	Abb63299 Drosophi1
43	291	6.9	2703	5	ABG70019	Abg70019 Larval vi
44	286	6.8	1390	5	ABU65152	Abu65152 Human NOV
45	286	6.8	1390	8	ADN61955	Adn61955 Human nov

RESULT 1

ADD89836
ID ADD89836 standard; protein, 813 AA.

AC ADD89836;

DT 29-JAN-2004 (first entry)

DE Human CAM-regulated phosphoprotein hTRAPP.

XX Human, hTRAPP; TARPP; neurodegenerative disease; Alzheimer's disease;

XX gene therapy; neuroprotective; nootropic.

XX Homo sapiens.

OS Key Location/Qualifiers

FT Misc-difference 68 /note= "Encoded by AGY"

XX WO2003083482-A1.

XX 09-OCT-2003.

XX 01-APR-2003; 2003WO-EP003364.

XX 02-APR-2002; 2002EP-00007522.

XX 02-APR-2002; 2002US-0368970P.

XX (EVOT-) EVOTEC NEUROSCIENCES GMBH.

XX Hipfler R, Hanes J, Von Der Kammer H, Pohlner J;

XX WPI; 2003-804100/75.

XX N-PSDB; ADD89837.

XX New nucleic acid useful for preparing a composition for treating,

XX preventing, diagnosing, prognosticating or monitoring the progression of

XX a neurodegenerative disease, e.g., Alzheimer's disease.

XX Claim 8; SEQ ID NO 1; 69pp; English.

XX The present sequence is the protein sequence of human CAM-regulated

XX phosphoprotein TRAPP (or hTRAPP), a previously undescribed human isoform

XX of murine CAM-regulated phosphoprotein ARPP-21. The invention discloses

XX the differential expression of hTRAPP in specific brain regions of

XX Alzheimer's disease patients. On the basis of this finding, a method is

provided for diagnosing or prognosticating a neurodegenerative disease, in particular Alzheimer's disease, or for determining whether a subject is at increased risk of developing such a disease. The gene encoding or hTARPP is used in therapeutic and prophylactic methods for treating or preventing Alzheimer's disease and related neurodegenerative disorders. A method of screening for modulating agents of neurodegenerative diseases is also provided. Transgenic animals comprising a non-native gene sequence coding for hTARPP are useful for screening, testing or validating compounds, agents or modulators, and developing diagnostics or therapeutics for treating neurodegenerative disease, particularly Alzheimer's disease. hTARPP protein is useful as a diagnostic target for detecting, or screening target for reagents or compounds for preventing, treating or ameliorating a neurodegenerative disease, preferably Alzheimer's disease.

Sequence 813 AA;

Query Match 100.0%; Score 4235; DB 7; Length 813;

Best Local Similarity 100.0%; Pred. No. 3.5e-278;

Matches 813; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
1 MSEGDLNQAIAEEGTBEETATPENGIVKSSLDDEEKLQRLAONORRRKSKGA 60
1 MSEGDLNQAIAEEGTBEETATPENGIVKSSLDDEEKLQRLAONORRRKSKGA 60
61 GKGLTRSLAVEESSARPGESLDOESIHQLSSFSLOEDSKRDKDSREKDKN 120
61 GKGLTRSLAVEESSARPGESLDOESIHQLSSFSLOEDSKRDKDSREKDKN 120
121 KDTSEKRIKMLSDCSQEYTDSTGIDHEFLINTLKNRDMILLKMEQIIDFIAD 180
121 KDTSEKRIKMLSDCSQEYTDSTGIDHEFLINTLKNRDMILLKMEQIIDFIAD 180
121 KDTSEKRIKMLSDCSQEYTDSTGIDHEFLINTLKNRDMILLKMEQIIDFIAD 180
121 KDTSEKRIKMLSDCSQEYTDSTGIDHEFLINTLKNRDMILLKMEQIIDFIAD 180
181 NNNHYKFPQSSSYRMLVHRVAAYFGLDHNDVQTKSYITIKTSSTRIPEDRFEHLKD 240
181 NNNHYKFPQSSSYRMLVHRVAAYFGLDHNDVQTKSYITIKTSSTRIPEDRFEHLKD 240
181 NNNHYKFPQSSSYRMLVHRVAAYFGLDHNDVQTKSYITIKTSSTRIPEDRFEHLKD 240
181 NNNHYKFPQSSSYRMLVHRVAAYFGLDHNDVQTKSYITIKTSSTRIPEDRFEHLKD 240
241 EKGEESORFLIKRNNSSIDKEDNOSVCSQESLFPENSLLDSDNINCEYTKRDLFRGN 300
241 EKGEESORFLIKRNNSSIDKEDNOSVCSQESLFPENSLLDSDNINCEYTKRDLFRGN 300
241 EKGEESORFLIKRNNSSIDKEDNOSVCSQESLFPENSLLDSDNINCEYTKRDLFRGN 300
241 EKGEESORFLIKRNNSSIDKEDNOSVCSQESLFPENSLLDSDNINCEYTKRDLFRGN 300
301 RDGSGRTSGRSSESENELKMSDHORAMSTDSNRLKPAWTKTASFGGITVLTNGD 360
301 RDGSGRTSGRSSESESENELKMSDHORAMSTDSNRLKPAWTKTASFGGITVLTNGD 360
301 RDGSGRTSGRSSESESENELKMSDHORAMSTDSNRLKPAWTKTASFGGITVLTNGD 360
301 RDGSGRTSGRSSESESENELKMSDHORAMSTDSNRLKPAWTKTASFGGITVLTNGD 360
361 STSSTRSTGKLSKAGSESSSSAGSSGSLSRTHPPLQSTPLVSGVAAGSPGCVPEENGIG 420
361 STSSTRSTGKLSKAGSESSSSAGSSGSLSRTHPPLQSTPLVSGVAAGSPGCVPEENGIG 420
361 STSSTRSTGKLSKAGSESSSSAGSSGSLSRTHPPLQSTPLVSGVAAGSPGCVPEENGIG 420
361 STSSTRSTGKLSKAGSESSSSAGSSGSLSRTHPPLQSTPLVSGVAAGSPGCVPEENGIG 420
421 GOVAASSTSYILLPLEATGTPRGSILLNPHTGOPFVNDGTPATYNPPTSGOPLRSAMV 480
421 GOVAASSTSYILLPLEATGTPRGSILLNPHTGOPFVNDGTPATYNPPTSGOPLRSAMV 480
421 GOVAASSTSYILLPLEATGTPRGSILLNPHTGOPFVNDGTPATYNPPTSGOPLRSAMV 480
421 GOVAASSTSYILLPLEATGTPRGSILLNPHTGOPFVNDGTPATYNPPTSGOPLRSAMV 480
481 GOSQOQPPQOQSPPOPOQVPPQPMAGPLVTQSVOGLQASQSSVOYAVAFPPQHLIP 540
481 GOSQOQPPQOQSPPOPOQVPPQPMAGPLVTQSVOGLQASQSSVOYAVAFPPQHLIP 540
481 GOSQOQPPQOQSPPOPOQVPPQPMAGPLVTQSVOGLQASQSSVOYAVAFPPQHLIP 540
481 GOSQOQPPQOQSPPOPOQVPPQPMAGPLVTQSVOGLQASQSSVOYAVAFPPQHLIP 540
541 VSPGTHFPMRDVATQFGMTLSRQSSGRTPEPPSPGVVPSLLMPQAPQPSYVASTGQ 600
541 VSPGTHFPMRDVATQFGMTLSRQSSGRTPEPPSPGVVPSLLMPQAPQPSYVASTGQ 600
541 VSPGTHFPMRDVATQFGMTLSRQSSGRTPEPPSPGVVPSLLMPQAPQPSYVASTGQ 600
541 VSPGTHFPMRDVATQFGMTLSRQSSGRTPEPPSPGVVPSLLMPQAPQPSYVASTGQ 600
601 QLPPTGFGSGGPPISQOVLQPPPSFGQFVQOQPPPMQPMPIYIPSGQYPISTIQQRMAP 660
601 QLPPTGFGSGGPPISQOVLQPPPSFGQFVQOQPPPMQPMPIYIPSGQYPISTIQQRMAP 660
601 QLPPTGFGSGGPPISQOVLQPPPSFGQFVQOQPPPMQPMPIYIPSGQYPISTIQQRMAP 660
601 QLPPTGFGSGGPPISQOVLQPPPSFGQFVQOQPPPMQPMPIYIPSGQYPISTIQQRMAP 660
661 VOYNARQSQOMQQAQAGYQPVLSGGQGFQGLIGVQOPPGSQNTYNNQGTPTVGSVWS 720
661 VOYNARQSQOMQQAQAGYQPVLSGGQGFQGLIGVQOPPGSQNTYNNQGTPTVGSVWS 720
661 VOYNARQSQOMQQAQAGYQPVLSGGQGFQGLIGVQOPPGSQNTYNNQGTPTVGSVWS 720
661 VOYNARQSQOMQQAQAGYQPVLSGGQGFQGLIGVQOPPGSQNTYNNQGTPTVGSVWS 720
721 YPTMSYQVPMYQSGQGLPQGSYQOPIMLPNOAGGSLPATGMPYTCNTPTPQNNRL 780
721 YPTMSYQVPMYQSGQGLPQGSYQOPIMLPNOAGGSLPATGMPYTCNTPTPQNNRL 780
721 YPTMSYQVPMYQSGQGLPQGSYQOPIMLPNOAGGSLPATGMPYTCNTPTPQNNRL 780
721 YPTMSYQVPMYQSGQGLPQGSYQOPIMLPNOAGGSLPATGMPYTCNTPTPQNNRL 780
```

781 IGPHCPSSTVPVMSASCRTNCASMSNAGQVKF 813
781 IGPHCPSSTVPVMSASCRTNCASMSNAGQVKF 813

RESULT 2
ADG17598
ADG17598 standard; protein; 813 AA.

ADG17598;

26-FEB-2004 (first entry)

Human Br137B protein amino acid sequence.

neuroprotective; ophthalmological; cytoskeletal; cardiact; antiarrhythmic;
gene therapy; molecular marker; drug target; detecting; diagnosing;
staging; monitoring; prognosticating; preventing; treating;
disease predisposition; abnormal gene expression; human;
neurological disorder; visual disorder; myopathy; heart failure;
arrhythmia; cancer; Br137B.

Homo sapiens.

WO2003085095-A2.

16-OCT-2003.

01-APR-2003; 2003WO-US009921.

01-APR-2002; 2002US-00112372.

24-MAY-2002; 2002US-0382614P.

13-JUN-2002; 2002US-00164717.

13-JUN-2002; 2002US-00167631.

24-JUN-2002; 2002US-0017917P.

30-JUL-2002; 2002US-0399125P.

(ORIG-) ORIGENE TECHNOLOGIES INC.

Jay G, Kovacs KF, Li X, Fan W, Shu Y, Yee A;

WPI; 2003-812725/76.

N-Psdb; ADG17597.

New expressed polynucleotides and polypeptides (e.g. OTB0949) useful as

molecular markers or as drug targets, for research, or for diagnosing,

preventing or treating diseases associated with abnormal gene expression

(e.g. cancer).

Claim 6; SEQ ID NO 48; 193pp; English.

This invention relates to a novel isolated DNA sequence and the proteins

encoded by them. The sequences disclosed may be useful during the

development of compounds with a neuroprotective, ophthalmological,

cytoskeletal, cardiact or antiarrhythmic activity. In addition the

sequences may be useful for gene therapy. Specifically claimed is an

isolated DNA comprising any of the 11 fully defined sequences of 1006-

7062 bp given in the specification. The polynucleotide encodes a

polypeptide comprising any of the 11 fully defined sequences of 89-1707

amino acids given in the specification. The DNA and protein are useful as

molecular markers, as drug targets, and for detecting, diagnosing,

staging, monitoring, prognosticating, preventing, treating or determining

predisposition to, various diseases and conditions associated with

abnormal expression of a gene in a subject (for example neurological or

visual disorders, myopathy, heart failure, arrhythmias or cancer). These

may also be used in research, drug discovery, clinical medicine or the

forensic science. The present sequence is that of a human protein of the

invention.

Sequence 813 AA;

Query Match 100.0%; Score 4235; DB 7; Length 813;

Best Local Similarity 100.0%; Pred. No. 3.5e-278;

1 ATGTCCTGAGCAAGGAGCCTGAATCAGGCAATAGCAGAGGAAGGAGGACTGAGCAGGAG 6C

DB 438 ATGTCTGACGACGAGACCTGATCAGCAATGACAGAGAGGAGGACTGACAGAG 497
QY 61 ACGGCCACTCCGAGAGACGGCATTTGTAATCAGAAAGTCTGGATGAAGAGAACTG 120
DB 498 ACGGCCACTCCGAGAGACGGCATTTGTAATCAGAAAGTCTGGATGAAGAGAACTG 557
QY 121 GAACGTCAGAGGCGGCTGGAGGCTCAGATCAAGAAAGAAATCCAAAGTCAGAGCA 180
DB 558 GAACGTCAGAGGCGGCTGGAGGCTCAGATCAAGAAAGAAATCCAAAGTCAGAGCA 617
QY 181 GGAAGAGTAACTGACTGCAAGCTTGTCTGTGAGGAATCTTCTGCCAGACCAAGA 240
DB 618 GGAAGAGTAACTGACTGCAAGCTTGTCTGTGAGGAATCTTCTGCCAGACCAAGA 677
QY 241 GGTGAAGTCTTCAGATCAG 261
DB 678 GGTGAAGTCTTCAGATCAG 698

RESULT 2
US-11-266-748A-147091/c
; Sequence 147091, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcription Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 147091
; LENGTH: 905
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (792)..(826)
; OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-147091

Query Match 10.7%; Score 260.6; DB 8; Length 905;
Best Local Similarity 99.6%; Pred. No. 2.2e-59;
Matches 260; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTCTGACGACGAGACCTGATCAGCAATGACAGAGAGGAGGACTGACAGAG 60
DB 468 ATGTCTGACGACGAGACCTGATCAGCAATGACAGAGAGGAGGACTGACAGAG 409
QY 61 ACGGCCACTCCGAGAGACGGCATTTGTAATCAGAAAGTCTGGATGAAGAGAACTG 120
DB 408 ACGGCCACTCCGAGAGACGGCATTTGTAATCAGAAAGTCTGGATGAAGAGAACTG 349
QY 121 GAACGTCAGAGGCGGCTGGAGGCTCAGATCAAGAAAGAAATCCAAAGTCAGAGCA 180

DB 348 GAACGTCAGAGGCGGCTGAGGCTCAGATCAAGAAAGAAATCCAAAGTCAGAGCA 289
QY 181 GGAAGAGTAACTGACTGCAAGCTTGTCTGTGAGGAATCTTCTGCCAGACCAAGA 240
DB 288 GGAAGAGTAACTGACTGCAAGCTTGTCTGTGAGGAATCTTCTGCCAGACCAAGA 229
QY 241 GGTGAAGTCTTCAGATCAG 261
DB 228 GGTGAAGTCTTCAGATCAG 208

RESULT 3
US-11-266-748A-381234
; Sequence 381234, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcription Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 381234
; LENGTH: 906
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-381234

Query Match 10.7%; Score 260.6; DB 8; Length 906;
Best Local Similarity 99.6%; Pred. No. 2.2e-59;
Matches 260; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTCTGACGACGAGACCTGATCAGCAATGACAGAGAGGAGGACTGACAGAG 60
DB 263 ATGTCTGACGACGAGACCTGATCAGCAATGACAGAGAGGAGGACTGACAGAG 322
QY 61 ACGGCCACTCCGAGAGACGGCATTTGTAATCAGAAAGTCTGGATGAAGAGAACTG 120
DB 323 ACGGCCACTCCGAGAGACGGCATTTGTAATCAGAAAGTCTGGATGAAGAGAACTG 382
QY 121 GAACGTCAGAGGCGGCTGGAGGCTCAGATCAAGAAAGAAATCCAAAGTCAGAGCA 180
DB 383 GAACGTCAGAGGCGGCTGGAGGCTCAGATCAAGAAAGAAATCCAAAGTCAGAGCA 442
QY 181 GGAAGAGTAACTGACTGCAAGCTTGTCTGTGAGGAATCTTCTGCCAGACCAAGA 240
DB 443 GGAAGAGTAACTGACTGCAAGCTTGTCTGTGAGGAATCTTCTGCCAGACCAAGA 502
QY 241 GGTGAAGTCTTCAGATCAG 261
DB 503 GGTGAAGTCTTCAGATCAG 523

No.	Score	Query	Match	Length	DB	ID	Description
1	262.2	10.7	2014	3	US-10-104-047-1766	Sequence 1766, App	
2	225.6	9.2	408	3	US-09-513-999C-22301	Sequence 22301, App	
3	75.6	3.1	7218	2	US-08-232-463-31	Sequence 14, App	
4	64.2	2.6	498	3	US-09-270-767-28545	Sequence 28545, App	
5	64.2	2.6	569	3	US-09-270-767-12730	Sequence 12730, App	
6	62.8	2.6	309	3	US-09-513-999C-12160	Sequence 12160, App	
7	53	2.2	592	3	US-09-385-219A-31	Sequence 31, App	
8	52	2.1	7218	2	US-08-232-463-31	Sequence 14, App	
9	50.8	2.1	2183	3	US-10-104-047-1064	Sequence 1064, App	
10	49.4	2.0	1359	4	US-09-297-648-4725	Sequence 4725, App	
11	46.8	1.9	2249	3	US-08-860-635A-18	Sequence 18, App	
12	46.8	1.9	2249	3	US-09-281-476-18	Sequence 18, App	
13	46.8	1.9	2249	3	US-09-910-087-18	Sequence 18, App	
14	46.6	1.9	289	3	US-09-207-005-17	Sequence 17, App	
15	46.6	1.9	289	3	US-09-244-996-17	Sequence 17, App	
16	46.4	1.9	2404	3	US-10-104-047-1740	Sequence 1740, App	
17	46.4	1.9	2462	3	US-09-922-364A-48	Sequence 48, App	
18	46.4	1.9	2462	3	US-09-254-590-48	Sequence 48, App	
19	46.4	1.9	2462	3	US-10-115-415-48	Sequence 48, App	
20	46.4	1.9	2462	3	US-10-116-360-48	Sequence 48, App	
21	46.4	1.9	2462	3	US-10-115-671-48	Sequence 48, App	
22	46.4	1.9	2462	3	US-10-115-695-48	Sequence 48, App	
23	46.4	1.9	2462	4	US-10-116-561-48	Sequence 48, App	

24	46.4	1.9	2462	5	US-10-115-688-48	Sequence 48, Appl
25	46.4	1.9	5082	5	US-09-497-822C-18	Sequence 18, Appl
26	46.4	1.9	4411529	3	US-09-103-860A-1	Sequence 1, Appl
27	46.2	1.9	612	3	US-09-902-554-157	Sequence 1357, Appl
28	46	1.9	623	3	US-09-043-30-5	Sequence 5, Appl1
29	46	1.9	1876	3	US-09-324-258-1	Sequence 1, Appl1
30	45.8	1.9	10348	2	US-08-457-273B-11	Sequence 41, Appl1
31	45.8	1.9	10348	3	US-08-556-419-13	Sequence 13, Appl
32	45.8	1.9	10348	3	US-09-041-886-14	Sequence 14, Appl1
33	45.8	1.9	10366	2	US-08-246-982A-5	Sequence 5, Appl1
34	45.8	1.9	10366	2	US-08-453-265-5	Sequence 5, Appl1
35	45.6	1.9	601	3	US-09-949-016-44973	Sequence 44973, Appl
36	45.6	1.9	601	3	US-09-949-016-44974	Sequence 44974, Appl
37	45.6	1.9	2514	3	US-09-949-016-1290	Sequence 1290, Appl
38	45.6	1.9	165651	3	US-09-949-016-11032	Sequence 11032, Appl
39	45.4	1.9	814	3	US-09-270-767-12898	Sequence 12898, Appl
40	45.2	1.9	255	3	US-09-488-039A-6195	Sequence 6195, Appl
41	45.2	1.9	3186	3	US-09-488-039A-5569	Sequence 5569, Appl
42	44.6	1.8	477	3	US-09-135-934-1	Sequence 1, Appl1
43	44.6	1.8	477	3	US-09-684-843A-1	Sequence 1, Appl1
44	44.2	1.8	1524	3	US-09-949-016-5791	Sequence 5791, Appl
45	44.2	1.8	2156	3	US-09-086-663A-78	Sequence 78, Appl

ALIGNMENTS

```

RESULT 1
US-10-104-047-1766
; Sequence 1766: Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; TITLE OF INVENTION: NO. 6943241e1 full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NOS: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIORITY APPLICATION NUMBER:
; PRIORITY FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1766
; LENGTH: 2014
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-1766

```

Query Match	Score	DB 3;	Length	2014
10.7%;	262.2;			

Best Local Similarity	98.9%	Pred.NO.	3,3e-63;
Matches	264;	Conservative	0;
Mismatches	3;	Indels	0;
Gaps	0;		

QY	2176	TCCTATCAGGTGCCAATGACCCAGGGTTCTCAAGACTGCCCCAGCAGTCAATCCAAAG	2235
Db	1227	TTTTTACAGGTCCAAATGACCCAGGGTTCTCAAGACTGCCCCAGCAGTCAATCCAAAG	1288
QY	2236	CCAATCATGTACTTAAACAAGCAGGTCAAGGGTCACTCCAGCCACTGGAATGCTGT	2295
Db	1287	CCAATCATGTACTTAAACAAGCAGGTCAAGGGTCACTCCAGCCACTGGAATGCTGT	1346
QY	2296	TACTGTATGTCAACCGGCCACCCCTCAAGAACAACTTAAAGCTGATTGGCCCCACACTGC	2355
Db	1347	TACTGTATGTCAACCGGCCACCCCTCAAGAACAACTTAAAGCTGATTGGCCCCACACTGC	1408
QY	2356	CCCTCCAGCACTGCCCAAGTATGCTAGCTGCAAGAACAAACTGTGCAAGATATGAGC	2415
Db	1407	CCCTCCAGCACTGTCCCAAGTATGCTAGCTGCAAGAACAAACTGTGCAAGATATGAGC	1466
QY	2416	AATGCTGGTTGGCAGGTCAAAATTTCTGA	2442
Db	1467	AATGCTGGTTGGCAGGTCAAAATTTCTGA	1493

```
US-09-513-999C-22301
; Sequence 22301, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclet, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59, US2, REG
; CURRENT APPLICATION NUMBER: US/09/513, 999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 22301
; LENGTH: 408
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 329
; OTHER INFORMATION: n=a, g, c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 330
; OTHER INFORMATION: d=a or g or t
US-09-513-999C-22301

Query Match
Best Local Similarity 9.2%; Score 225.6; DB 3; Length 408;
Matches 225; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 36 AGAGAGAGAGGAGTCTGACGAGAGACGCGCATCTCCAGAGAAAGCGCATTTGTTAAATCAGA 95
DB 1 AGAGAGAGAGGAGTCTGACGAGAGACGCGCATCTCCAGAGAAAGCGCATTTGTTAAATCAGA 60
QY 96 AAGCTGTGATGAAGGAGGAAAGTGAAGTCTGACGAGGGGGCTGAGGGCTCAGAAATCAGA 155
DB 61 AAGCTGTGATGAAGGAGGAAAGTGAAGTCTGACGAGGGGGCTGAGGGCTCAGAAATCAGA 120
QY 156 AAGAGAGAAATTCAGTCAAGTCAAGGAGCAGAGAAAGTAACTGACTCGCAGCTTGTCTGTG 215
DB 121 AAGAGAGAAATTCAGTCAAGTCAAGGAGCAGAGAAAGTAACTGACTCGCAGCTTGTCTGTG 180
QY 216 TGAGGAATCTTCTGCGCAGACCGAGAGGTGAAGTCTTCAGATCAG 261
DB 181 TGAGGAATCTTCTGCGCAGACCGAGAGGTGAAGTCTTCAGATCAG 226

RESULT 3
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
```

```
US-09-513-999C-22301
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-9300
; TELEFAX: (703) 683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-Fls
US-08-232-463-14

Query Match
Best Local Similarity 3.1%; Score 75.6; DB 2; Length 7218;
Matches 12; Conservative 241; Mismatches 135; Indels 0; Gaps 0;

QY 7 GAGCAAGAGACCTGGAATCAGCAATGACAGAGAGAGGAGTACGACAGAGACGCGC 66
DB 1431 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1372
QY 67 ACTCCAGAGAACGCGCATTTGTTAAATCAGAAAGTCTGATGAGAGAGAGAACTGCACTG 126
DB 1371 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1312
QY 127 CAGAGGGGGCTGAGGGCTCAGAAATCAGAGAAAGAAATCCAAAGTCAAGAGCAGAGAA 186
DB 1311 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1252
QY 187 GGTAACTGACTCCAGACTTGTCTGTGAGAAATCTTCTGCGCAGACGAGAGTGA 246
DB 1251 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1192
QY 247 AGCTTCAGATCAGAAATCAATTCATTACAGCTTTTCCAGCTTTCAGCCTGCAAG 306
DB 1191 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1132
QY 307 GAGGATAAATCTAGAGAAAGATGACTGTGAAGAGAAAGAAAGATTAACAAGAT 366
DB 1131 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1072
QY 367 AAAACCTCTGAAAAACCCAGATCAGAA 394
DB 1071 RRRRRATGCAAGCTCCTGACCTGCA 1044

RESULT 4
US-09-270-767-28545/c
; Sequence 28545, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 28545
```

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 5, 2006, 09:44:00 ; Search time 11518 Seconds

(without alignments)
11855.792 Million cell updates/sec

Title: US-10-509-950-2

Perfect score: 2441.6
Sequence: 1 agtccgagcagagagacc.....gtcgcagcagcraactctga 2442

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 48236798 seqs, 27959665780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Matching first 45 summaries

Database :

EST
1: gb_est1:*
2: gb_est3:*
3: gb_est4:*
4: gb_est5:*
5: gb_est6:*
6: gb_est7:*
7: gb_est8:*
8: gb_est9:*
9: gb_est10:*
10: gb_est11:*
11: gb_est12:*
12: gb_est13:*
13: gb_est14:*
14: gb_est15:*

Pred. No. is the number of results predicted by a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1944	79.6	2505	6	HSMB01390
2	1815.8	74.4	2914	6	AK034376
3	1724.8	70.6	3180	6	AK032417
4	1641.2	67.2	2991	6	AK032543
5	938	38.4	2806	6	AK137974
6	910.4	37.3	1550	8	CO635664
7	757.6	31.0	845	2	B1553063
8	748	30.6	925	10	DV909010
9	743	30.4	1053	3	BM926937
10	742.6	30.4	929	10	DV909014
11	695	28.5	947	10	DV879970
12	662.4	27.1	743	9	DN991171
13	641	26.3	991	4	BY761855
14	641	26.3	991	4	BY761855
15	630	25.8	1864	6	AK029467
16	620.2	25.4	922	6	AK153892
17	615	25.2	749	4	BU164142
18	605.8	24.8	609	9	DA529811
19	600.4	24.6	1056	10	DV797399

20	590.8	24.2	866	4	EX461002
21	588.8	24.1	753	5	CJ464930
22	586	24.0	782	4	BX431175
23	571.4	23.4	574	9	DA487336
24	571	23.4	571	9	DA494904
25	569.6	23.3	740	5	CJ439713
26	565.6	23.2	837	3	BQ217730
27	562.8	23.1	926	3	BP160144
28	552.8	22.6	634	9	DN427693
29	551.2	22.6	629	5	CK003565
30	549.4	22.5	563	9	DA764999
31	547.8	22.4	575	9	DA179590
32	540.2	22.1	654	7	BB626408
33	540.2	22.1	891	3	BP160316
34	538.2	22.0	1946	6	AK045507
35	536.6	22.0	711	3	BP162992
36	534.6	21.9	1688	6	AK050613
37	534.4	21.9	765	3	BO179621
38	526.4	21.6	728	3	BM951156
39	516	21.1	566	9	DA260904
40	516	21.1	595	9	DB172765
41	511.8	21.0	806	5	CK460159
42	506.4	20.7	728	3	BU342596
43	506.4	20.7	800	3	BU604247
44	493	20.2	575	9	DA389805
45	491.2	20.1	640	7	BB622926

ALIGNMENTS

RESULT 1
LOCUS HSM801390
DEFINITION Homo sapiens mRNA; cDNA DKFP566N1047 (from clone DKFP566N1047).
ACCESSION AL133109
VERSION AL133109.1 GI:6453595
KEYWORDS
SOURCE HTC.
ORGANISM Homo sapiens (human)
Homo sapiens
Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 2505)
Ottewaelder, B., Obermaier, B., Deutschenbaun, S., Schaipp, A.,
Mewes, H.W., Weill, B., Amid, C., Oeanger, A., Fobo, G., Han, M., and
Wiemann, S.
The German cDNA Consortium
Direct Submission
Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequencing by Medigenomix (Martinsried/Germany) within the cDNA
This clone (DKFP566N1047) is available at the RZPD Deutsches
Resourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFP566N1047
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/.

FEATURES

source
1..2505
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKFP566N1047"
/db_xref="taxon:9606"
/clone="DKFP566N1047"
/issue_type="kidney"
/clone_lib="566 (synonym: hfkcd2). Vector pAMP1, host
XL-2Bp1; sites NciI + SalI"
/dev_stage="fetal"
/note="cyclic AMP-regulated phosphoprotein, 21 kD."

gene N-terminus truncated, differentially spliced"
1..2505
/gene="DKFZp566N1047"
<3..1958
/gene="DKFZp566N1047"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAB61414.1"
/db_xref="GI:6453596"
/db_xref="GOA:O9UF93"
/db_xref="UniProtKB/TrEMBL:O9UF93"
/translation="RDRMLLKMOEIIIDTADNNHYKKPEPOMSSYORMLVHVAAY
FGLDHVDQTKGVIIINKSTSTRIPROFCEHLKDEKESORFLIKRNSIIDKED
NOSVCSQESLFEVNRILEDNINCHETKTKKQLRGNFDSGSRISGSSSENLKRS
DHQWMSSTDSNENRLKPKMTKTSFGIIVLTRDSTSTRSTKSGSESS
SAGSSGLSTRHPLQSTPLVSGVAAGSPCVPEPENGIGQVAPSTSYILPLEA
TGIPSGIILNPHGTQGFVNPDGPALYNPTSQPLRSAMVQSQPQOQPSPP
QOQVOPQPMAGPLVTSVOGLQASSQSVQYPAVSFPPOHLFVSTQHPMRDVA
TORGMILSFOSGGEPEPSGPVPSLSLMDPAPQOSQVIASTGQQLPTGFSGSGP
PIQOYIQQPPSPQGFYQPPPAQMPYRYSQYPISTTQYRPMAPVYNQSRQ
MPOAQDAGTQPVLSGQGFQGLIGVQDQPSQVYNQGTVPQSVMSVPTMSSTQ
VPMTOGSGGLPQOYSTQPIMLPNOAGGSLPATGMVPCNVTPPTPQNNRLIGPHCP
STVPSVASCRITCASMSNAGWQVKE"

ORIGIN

Query Match 79.6%; Score 1944; DB 6; Length 2505;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1957; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

483 CAGGACAGAGATGATCTTTGAAAATGAGCAGAAATTAATGATTTCAATGCTGACAA 542.
DB 2 CAGGACAGAGATGATCTTTGAAAATGAGCAGAAATTAATGATTTCAATGCTGACAA 61
543 CAATATCATTAATAAAGTTCCCTCAGATGTCATGATCAGAGATGCTGTGATCG 602
DB 62 CAATATCATTAATAAAGTTCCCTCAGATGTCATGATCAGAGATGCTGTGATCG 121
603 AGTGGCAGCTTATTTGGATTGATCAATGTGATCAAAAGAAAATCTGTATCAT 662
DB 122 AGTGGCAGCTTATTTGGATTGATCAATGTGATCAAAAGAAAATCTGTATCAT 181
663 CAACAAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 722
DB 182 CAACAAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 241
723 AAAAGTGAAGATCCAGAAAGCGTTTATCTTGAAGCAGATTAATCTAGATATGATA 782
DB 242 AAAAGTGAAGATCCAGAAAGCGTTTATCTTGAAGCAGATTAATCTAGATATGATA 301
783 AGAAGACATCATGATGCTGCTCCAGAAAGCCTTTTGTGAAAACAGTAGGCTCTT 842
DB 302 AGAAGACATCATGATGCTGCTCCAGAAAGCCTTTTGTGAAAAC--AGGCTCTT 358
843 GGAAGACATTAATGATGAGCCTATTAAGAAAAGCAGCTCTTTGGGGGCAACAG 902
DB 359 GGAAGACATTAATGATGAGCCTATTAAGAAAAGCAGCTCTTTGGGGGCAACAG 418
903 AGATGCTCAGAGGAAATCTGTGGAGTCCGACAGAGAGCTCAGAAAATGATCAAGTG 962
DB 419 AGATGCTCAGAGGAAATCTGTGGAGTCCGACAGAGAGCTCAGAAAATGATCAAGTG 478
963 GTCTGACCAACAAAGGGGCTGAGACAGACAGATCTCCGACAGTTCCAAACCGCAATCTAAA 1022
DB 479 GTCTGACCAACAAAGGGGCTGAGACAGACAGATCTCCGACAGTTCCAAACCGCAATCTAAA 538
1023 GCCCGCATGACCAAGAGCGGAGTTTGGGGGATCATCAGGTGCTGACCAAGGGGTGACAG 1082
DB 539 GCCCGCATGACCAAGAGCGGAGTTTGGGGGATCATCAGGTGCTGACCAAGGGGTGACAG 598
1083 CACTTCCAGTACTAGAGTACCGGAAAGCTGTCCAAAGCAGGTTCCAGTCTTCCAGACAG 1142
DB 599 CACTTCCAGTACTAGAGTACCGGAAAGCTGTCCAAAGCAGGTTCCAGTCTTCCAGACAG 658

1143 TGCAGGCTCCTCAGATTCGCTGTCCGCAACCCATCCACTCTCCAGACACACCCCTAGT 1202
DB 659 TGCAGGCTCCTCAGATTCGCTGTCCGCAACCCATCCACTCTCCAGACACACCCCTAGT 718
1203 CTCAGGTGTGGCAGCTGTGCTCTCCAGAGCTGTGTGCTTATCCAGAGATGAAATAGGGGG 1262
DB 719 CTCAGGTGTGGCAGCTGTGCTCTCCAGAGCTGTGTGCTTATCCAGAGATGAAATAGGGGG 778
1263 CCAAGTTGCTCCAGACAGACACAGCTCATCTCTCTTCCACTTGAAGCTGCAAGGCAT 1322
DB 779 CCAAGTTGCTCCAGACAGACACAGCTCATCTCTCTTCCACTTGAAGCTGCAAGGCAT 838
1323 CCGGCTGGAAGATCCTTTTATCCACACAGAGGACCCCTTGTGAATCCCGATGG 1382
DB 839 CCGGCTGGAAGATCCTTTTATCCACACAGAGGACCCCTTGTGAATCCCGATGG 898
1383 AACTCTGCAATATTAACACCCACCCACAGTACAGAGCCCTGGAAGCCCATGTGGG 1442
DB 899 AACTCTGCAATATTAACACCCACCCACAGTACAGAGCCCTGGAAGCCCATGTGGG 958
1443 GCAGTCCCAACAGACCGGCTCAACAGACAGCCCTCCCGCAGACCCCAACAGAGTCCA 1502
DB 959 GCAGTCCCAACAGACCGGCTCAACAGACAGCCCTCCCGCAGACCCCAACAGAGTCCA 1018
1503 GCCACGCGACGACAGATGAGGAGCCCTGTGATCATCATGCTGTCCAGGGGCTGACAGG 1562
DB 1019 GCCACGCGACGACAGATGAGGAGCCCTGTGATCATCATGCTGTCCAGGGGCTGACAGG 1078
1563 TTCCTCCAGTACAGTCAATATCCGGCAGTCTCTTTTCTCCCGACAGCTCTCACTGT 1622
DB 1079 TTCCTCCAGTACAGTCAATATCCGGCAGTCTCTTTTCTCCCGACAGCTCTCACTGT 1138
1623 GTCTCCACGACAGACCTTTCCCATGAGATGATGTGGCAACAAGTTTGGCCAGATGAC 1682
DB 1139 GTCTCCACGACAGACCTTTCCCATGAGATGATGTGGCAACAAGTTTGGCCAGATGAC 1198
1683 CTTGAGCGGACGATCTCGGGGAGAGCTCTGAAACCCCATCAGGTCCTGTCACTCCATC 1742
DB 1199 CTTGAGCGGACGATCTCGGGGAGAGCTCTGAAACCCCATCAGGTCCTGTCACTCCATC 1258
1743 CTCCTTATGACACAGCGGAGCCAGACAGCCAGCTATGATGTGCTCTCAACGCGACAG 1802
DB 1259 CTCCTTATGACACAGCGGAGCCAGACAGCCAGCTATGATGTGCTCTCAACGCGACAG 1318
1803 GCTTCTCAAGAGGATTCAGAGCTGTGGCCTCTCCATCTCCAGAGAGTCTCTCCAGCC 1862
DB 1319 GCTTCTCAAGAGGATTCAGAGCTGTGGCCTCTCCATCTCCAGAGAGTCTCTCCAGCC 1378
1863 CCGTCCCTCAACAGAGGATTCGTCAAACAGCCTCCGCTGACAGATGCTGTATATTA 1922
DB 1379 CCGTCCCTCAACAGAGGATTCGTCAAACAGCCTCCGCTGACAGATGCTGTATATTA 1438
1923 TTACCATCTGTAGTACGATACCTCACTCAACACAGCAACAGTACCGGCCATGCCCCGGT 1982
DB 1439 TTACCATCTGTAGTACGATACCTCACTCAACACAGCAACAGTACCGGCCATGCCCCGGT 1498
1983 TCAGTACACGCTCAAGAGAGTCAACAGATGCCACAGGACGACAGGAAGAGGTTTACCA 2042
DB 1499 TCAGTACACGCTCAAGAGAGTCAACAGATGCCACAGGACGACAGGAAGAGGTTTACCA 1558
2043 GCCAGCTTGTCTGTCAACAGGATTCGTCAAACAGCCTCAATAGAGTGCAGACGACCTCA 2102
DB 1559 GCCAGCTTGTCTGTCAACAGGATTCGTCAAACAGCCTCAATAGAGTGCAGACGACCTCA 1618
2103 GAGTCAAGACGTATTAATTAACCAACAGGAATCTCGGTGCAAAAGCGTATGTTTCTTA 2162
DB 1619 GAGTCAAGACGTATTAATTAACCAACAGGAATCTCGGTGCAAAAGCGTATGTTTCTTA 1678
2163 CCCAACAATGCTCTTATCAGGAGCCAAATGACCAAGGGTTTCCAAAGAGTGCAGGACGAC 2222
DB 1679 CCCAACAATGCTCTTATCAGGAGCCAAATGACCAAGGGTTTCCAAAGAGTGCAGGACGAC 1738
2223 GTCATACCAACAGCAATCATGTACTTAACAGGACGAGTCAAGGGTCACTCCAGGACAC 2282

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 5, 2006, 09:44:00 ; Search time 1406 Seconds
(without alignments)

12109.695 Million cell updates/sec

Title: US-10-509-950-2

Perfect score: 2441.6
Sequence: 1 atgtctgagcaagagagacct.....gttcgacagtcacattctga 2442

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

1: N_Geneseq_B: *
2: geneeqn1980s: *
3: geneeqn1990s: *
4: geneeqn2000s: *
5: geneeqn2001as: *
6: geneeqn2001bs: *
7: geneeqn2002as: *
8: geneeqn2002bs: *
9: geneeqn2003as: *
10: geneeqn2003bs: *
11: geneeqn2003cs: *
12: geneeqn2004as: *
13: geneeqn2004bs: *
14: geneeqn2005s: *
15: geneeqn2006s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2441.6	100.0	3212	10	ADD89837 Human CAM
2	2441.2	100.0	3332	10	ADG17597 Human Brl
3	2324.8	95.2	3369	10	ADG17593 Human TAR
4	2311.2	94.7	3272	10	ADG17599 Human Brl
5	2287.2	93.7	3374	10	ADG17595 Human Brl
6	2200.8	90.1	2813	6	ABQ60920 Human Brl
7	1746.4	71.5	2548	5	AA662277 DKFZp566n
8	1366.2	56.0	1942	5	AA662275 DNA encod
9	750.2	30.7	1305	5	AA662275 DNA encod
10	650.4	26.6	1488	5	AA662275 DNA encod
11	650.4	26.6	1488	5	AA662275 DNA encod
12	413.2	16.5	468	5	ABAI3329 Human ner
13	362.4	14.8	4340	13	ADRU7590 Full leng
14	353.4	14.5	625	5	ABAI2203 Human ner
15	308.8	12.6	4330	8	ACCT2057 BCU0147B
16	308.8	12.6	4331	10	ADG37198 Nuclear f
17	274.8	11.3	3839	8	ACCS9933 Human IRA
18	274.8	11.3	4373	10	ADG37200 Nuclear f

19	262.2	10.7	2014	10	ADB63612	ADB63612 Human CDN
20	260.6	10.7	352	6	ABA04824	ABA04824 Human hkr
21	260.6	10.7	1006	10	ADG17601	ADG17601 Human Brl
22	260.6	10.7	1048	10	ADD29568	ADD29568 Human tum
23	260.6	10.7	4276	8	ACCT2056	ACCT2056 BCU0147A
24	259	10.6	876	3	AAA47416	AAA47416 Sequence
25	257.4	10.5	371	2	AAQ60619	AAQ60619 Human bra
26	257.4	10.5	371	5	AA662273	AA662273 DNA encod
27	257.4	10.5	371	5	AA662273	AA662273 DNA encod
28	255.8	10.5	4272	15	AEF74808	AEF74808 Human pol
29	233.8	9.6	579	12	ACH71408	ACH71408 Human gen
30	231.2	9.5	274	12	ACH85164	ACH85164 Human gen
31	229.4	9.4	567	5	AA662274	AA662274 DNA encod
32	225.6	9.2	408	3	AA662274	AA662274 DNA encod
33	207.2	8.5	308	5	AA662274	AA662274 DNA encod
34	207.2	8.5	308	5	AA662274	AA662274 DNA encod
35	207.2	8.5	308	5	AA662274	AA662274 DNA encod
36	206.6	8.5	646	13	ADQ52802	ADQ52802 Novel can
37	203.2	8.3	1305	5	AA662275	AA662275 DNA encod
38	203.2	8.3	1942	5	AA662275	AA662275 DNA encod
39	177	7.2	404	9	ACH46873	ACH46873 Human inf
40	160	6.6	1030	10	ADBS3792	ADBS3792 Primary r
41	160	6.6	1030	10	ADBS3792	ADBS3792 Primary r
42	150.8	6.2	285	6	ABN77809	ABN77809 Human ORF
43	136.2	5.6	2906	10	ADG37196	ADG37196 Nuclear f
44	134.8	5.5	510	12	ACH71623	ACH71623 Human gen
45	129	5.3	302	12	ACH85379	ACH85379 Human gen

ALIGNMENTS

RESULT 1	ADD89837 standard; cDNA; 3212 BP.
ID	ADD89837
AC	ADD89837
DT	29-JAN-2004 (first entry)
DE	Human cAMP-regulated phosphoprotein hTARPP cDNA.
KW	Human; hTARPP; TARPP; neurodegenerative disease; Alzheimer's disease;
KW	gene therapy; neuroprotective; nootropic; transgenic; gene; ss.
OS	Homo sapiens.
PH	Key
PH	Location/Qualifiers
FT	CDS
FT	240..2681
FT	/*tag= a
FT	/product= "hTARPP"
FT	/note= "The CDS is also claimed in Claim 3"
XX	WO2003083482-A1.
XX	09-OCT-2003.
XX	01-APR-2003; 2003WO-EP003364.
XX	02-APR-2002; 2002EP-00007522.
XX	02-APR-2002; 2002US-0368970P.
XX	(EVOT-) EVOTEC NEUROSCIENCES GMBH.
XX	Hipfler R, Hanes J, Von Der Kammer H, Pohlner J;
XX	WPI; 2003-804100/75.
XX	P-PSDB; ADD89836.
XX	New nucleic acid useful for preparing a composition for treating,
XX	preventing, diagnosing, prognosticating or monitoring the progression of
XX	a neurodegenerative disease, e.g. Alzheimer's disease.

PS Claim 3; SEQ ID NO 3; 69pp; English.

XX The present sequence, the coding region of which is also claimed, is that
 CC of cDNA encoding human CAMP-regulated phosphoprotein TRAP (or hrARP), a
 CC previously undescribed human isoform of murine CAMP regulated
 CC phosphoprotein ARP-21. The cDNA comprises an alignment of human
 CC expressed sequence tag clones that were obtained from a database
 CC screening using a differentially expressed human cDNA fragment AD98938.
 CC The invention discloses the differential expression of hrARP in specific
 CC brain regions of Alzheimer's disease patients. On the basis of this
 CC finding, a method is provided for diagnosing or prognosticating a
 CC neurodegenerative disease, in particular Alzheimer's disease, or for
 CC determining whether a subject is at increased risk of developing such a
 CC disease. The gene encoding hrARP is used in therapeutic and prophylactic
 CC methods for treating or preventing Alzheimer's disease and related
 CC neurodegenerative disorders. A method of screening for modulating agents
 CC of neurodegenerative diseases is also provided. Transgenic animals
 CC comprising a non-native gene sequence coding for hrARP are useful for
 CC screening, testing or validating compounds, agents or modulators, and
 CC developing diagnostics or therapeutics for treating neurodegenerative
 CC disease, particularly Alzheimer's disease. The encoded hrARP protein is
 CC useful as a diagnostic target for detecting, or screening target for
 CC reagents or compounds for preventing, treating or ameliorating a
 CC neurodegenerative disease, preferably Alzheimer's disease.

XX Sequence 3212 BP; 938 A; 795 C; 745 G; 733 T; 0 U; 1 Other;

Query Match 100.0%; Score 2441.6; DB 10; Length 3212;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTGAGCAAGAGACCTGATCAGGCAATGAGAGAGAGGAGCTAGAGAGAG 60
 Db ATGCTGAGCAAGAGACCTGATCAGGCAATGAGAGAGAGGAGCTAGAGAGAG 299
 QY 61 ACGGCACTCCAGAGACCGCATTTGTTAATCAGAAAGTCTGATGAGAGAGAACTG 120
 Db ACGGCACTCCAGAGACCGCATTTGTTAATCAGAAAGTCTGATGAGAGAGAACTG 359
 QY 121 GAATGAGAGAGCGGCTGAGAGCTCAGAAATCAAGAAAGAAATCCAGTCAAGAGCA 180
 Db GAATGAGAGAGCGGCTGAGAGCTCAGAAATCAAGAAAGAAATCCAGTCAAGAGCA 419
 QY 360 GAATGAGAGAGCGGCTGAGAGCTCAGAAATCAAGAAAGAAATCCAGTCAAGAGCA 419
 Db GAATGAGAGAGCGGCTGAGAGCTCAGAAATCAAGAAAGAAATCCAGTCAAGAGCA 419
 QY 181 GGAAGAGGTAACTGACTGCAAGCTTGTCTGAGAGAAATCTTCTGCAAGACAGCA 240
 Db GGAAGAGGTAACTGACTGCAAGCTTGTCTGAGAGAAATCTTCTGCAAGACAGCA 479
 QY 420 GGAAGAGGTAACTGACTGCAAGCTTGTCTGAGAGAAATCTTCTGCAAGACAGCA 479
 Db GGAAGAGGTAACTGACTGCAAGCTTGTCTGAGAGAAATCTTCTGCAAGACAGCA 479
 QY 241 GGTGAAAGTCTCAGGATCAGAAATCAATTTCACTTCCAGTTTTCAGACCTG 300
 Db GGTGAAAGTCTCAGGATCAGAAATCAATTTCACTTCCAGTTTTCAGACCTG 539
 QY 480 GGTGAAAGTCTCAGGATCAGAAATCAATTTCACTTCCAGTTTTCAGACCTG 539
 Db GGTGAAAGTCTCAGGATCAGAAATCAATTTCACTTCCAGTTTTCAGACCTG 539
 QY 301 CAAAGAGAGGATTAATCTTGAAGAAAGTACTCTGAAAGAGAAAGAAAGATTAAGC 360
 Db CAAAGAGAGGATTAATCTTGAAGAAAGTACTCTGAAAGAGAAAGAAAGATTAAGC 599
 QY 540 CAAAGAGAGGATTAATCTTGAAGAAAGTACTCTGAAAGAGAAAGAAAGATTAAGC 599
 Db CAAAGAGAGGATTAATCTTGAAGAAAGTACTCTGAAAGAGAAAGAAAGATTAAGC 599
 QY 361 AAGAGTAAACCTCTGAAAGAACCAAGATCAGAAATCTTCAAAAGATGAGCCAGAA 420
 Db AAGAGTAAACCTCTGAAAGAACCAAGATCAGAAATCTTCAAAAGATGAGCCAGAA 659
 QY 600 AAGAGTAAACCTCTGAAAGAACCAAGATCAGAAATCTTCAAAAGATGAGCCAGAA 659
 Db AAGAGTAAACCTCTGAAAGAACCAAGATCAGAAATCTTCAAAAGATGAGCCAGAA 659
 QY 421 TACAGGATTTCTACAGGATGAGTCAAGATTTCTGATTAACATTAAGAAATAT 480
 Db TACAGGATTTCTACAGGATGAGTCAAGATTTCTGATTAACATTAAGAAATAT 719
 QY 660 TACAGGATTTCTACAGGATGAGTCAAGATTTCTGATTAACATTAAGAAATAT 719
 Db TACAGGATTTCTACAGGATGAGTCAAGATTTCTGATTAACATTAAGAAATAT 719
 QY 481 TCCAGGAGAGAGATGATCTTTTGAAGAAAGAGAGAGAAATTAATTTCAATGCTGAC 540
 Db TCCAGGAGAGAGATGATCTTTTGAAGAAAGAGAGAGAAATTAATTTCAATGCTGAC 779
 QY 720 TCCAGGAGAGAGATGATCTTTTGAAGAAAGAGAGAGAAATTAATTTCAATGCTGAC 779
 Db TCCAGGAGAGAGATGATCTTTTGAAGAAAGAGAGAGAAATTAATTTCAATGCTGAC 779
 QY 541 AACATTAATCATTAATAAAGTCTCCTCAGATGTCATGATCAGAGAGATGCTTGCAT 600
 Db AACATTAATCATTAATAAAGTCTCCTCAGATGTCATGATCAGAGAGATGCTTGCAT 839
 QY 780 AACATTAATCATTAATAAAGTCTCCTCAGATGTCATGATCAGAGAGATGCTTGCAT 839
 Db AACATTAATCATTAATAAAGTCTCCTCAGATGTCATGATCAGAGAGATGCTTGCAT 839
 QY 601 CGAGTGGAGCTTATTTTGGATTTGATCAATGTGGATCAACAGAGAAATCTGTATC 660
 Db CGAGTGGAGCTTATTTTGGATTTGATCAATGTGGATCAACAGAGAAATCTGTATC 660

Db CGAGTGGAGCTTATTTTGGATTTGATCAATGTGGATCAACAGAGAAATCTGTATC 899
 QY 661 ATCAACAAGACAGAGAGACCAAGATTAACAGAGCAAGGTTTGGACATTTAAAGAT 720
 Db ATCAACAAGACAGAGAGACCAAGATTAACAGAGCAAGGTTTGGACATTTAAAGAT 959
 QY 900 ATCAACAAGACAGAGAGACCAAGATTAACAGAGCAAGGTTTGGACATTTAAAGAT 959
 Db ATCAACAAGACAGAGAGACCAAGATTAACAGAGCAAGGTTTGGACATTTAAAGAT 959
 QY 721 GAAAAGGTGAAGATCCAGAGAGGTTTATCTTGAAGAGAGATTAAGTATGAT 780
 Db GAAAAGGTGAAGATCCAGAGAGGTTTATCTTGAAGAGAGATTAAGTATGAT 1019
 QY 960 GAAAAGGTGAAGATCCAGAGAGGTTTATCTTGAAGAGAGATTAAGTATGAT 1019
 Db GAAAAGGTGAAGATCCAGAGAGGTTTATCTTGAAGAGAGATTAAGTATGAT 1019
 QY 781 AAGAGAGCAATCAGTCACTTTGCTCCAGAGAAAGCTTTTGTGAAACAGTGGCTC 840
 Db AAGAGAGCAATCAGTCACTTTGCTCCAGAGAAAGCTTTTGTGAAACAGTGGCTC 1079
 QY 1020 AAGAGAGCAATCAGTCACTTTGCTCCAGAGAAAGCTTTTGTGAAACAGTGGCTC 1079
 Db AAGAGAGCAATCAGTCACTTTGCTCCAGAGAAAGCTTTTGTGAAACAGTGGCTC 1079
 QY 841 TTGAGAGCAATTAATATGATGAGAGCACTTAAGAAAGAGAGCTTTTGGGGCAAC 900
 Db TTGAGAGCAATTAATATGATGAGAGCACTTAAGAAAGAGAGCTTTTGGGGCAAC 1139
 QY 1080 TTGAGAGCAATTAATATGATGAGAGCACTTAAGAAAGAGAGCTTTTGGGGCAAC 1139
 Db TTGAGAGCAATTAATATGATGAGAGCACTTAAGAAAGAGAGCTTTTGGGGCAAC 1139
 QY 901 AGAGATGCTCAGAGAGAGCAATCTGGAGAGTGCAGAGAGAGCTCAAGAAATGACTCAG 960
 Db AGAGATGCTCAGAGAGAGCAATCTGGAGAGTGCAGAGAGAGCTCAAGAAATGACTCAG 1199
 QY 1140 AGAGATGCTCAGAGAGAGCAATCTGGAGAGTGCAGAGAGAGCTCAAGAAATGACTCAG 1199
 Db AGAGATGCTCAGAGAGAGCAATCTGGAGAGTGCAGAGAGAGCTCAAGAAATGACTCAG 1199
 QY 961 TGTCTGACCAACCAAGGCTCTGAGAGAGCAAGCTCCAGAGTTTCCAGCCGATCTA 1020
 Db TGTCTGACCAACCAAGGCTCTGAGAGAGCAAGCTCCAGAGTTTCCAGCCGATCTA 1259
 QY 1200 TGTCTGACCAACCAAGGCTCTGAGAGAGCAAGCTCCAGAGTTTCCAGCCGATCTA 1259
 Db TGTCTGACCAACCAAGGCTCTGAGAGAGCAAGCTCCAGAGTTTCCAGCCGATCTA 1259
 QY 1021 AAGCCCGCATGACCAACAGGCGAGTTTGGGGGCTATCAGGAGTGCAGAGGGGTGAC 1080
 Db AAGCCCGCATGACCAACAGGCGAGTTTGGGGGCTATCAGGAGTGCAGAGGGGTGAC 1319
 QY 1260 AAGCCCGCATGACCAACAGGCGAGTTTGGGGGCTATCAGGAGTGCAGAGGGGTGAC 1319
 Db AAGCCCGCATGACCAACAGGCGAGTTTGGGGGCTATCAGGAGTGCAGAGGGGTGAC 1319
 QY 1081 AGCACTTCAGTACTAGAGAGTACCGGAGAGCTGTCGAAAGAGTTCGAGTCTTCCAGC 1140
 Db AGCACTTCAGTACTAGAGAGTACCGGAGAGCTGTCGAAAGAGTTCGAGTCTTCCAGC 1379
 QY 1320 AGCACTTCAGTACTAGAGAGTACCGGAGAGCTGTCGAAAGAGTTCGAGTCTTCCAGC 1379
 Db AGCACTTCAGTACTAGAGAGTACCGGAGAGCTGTCGAAAGAGTTCGAGTCTTCCAGC 1379
 QY 1141 AGTCAAGCTCTCTCAGATTCGCTGCTCCGACCCATCCAGCTCTCTCAGAGCAACCCCTA 1200
 Db AGTCAAGCTCTCTCAGATTCGCTGCTCCGACCCATCCAGCTCTCTCAGAGCAACCCCTA 1439
 QY 1380 AGTCAAGCTCTCTCAGATTCGCTGCTCCGACCCATCCAGCTCTCTCAGAGCAACCCCTA 1439
 Db AGTCAAGCTCTCTCAGATTCGCTGCTCCGACCCATCCAGCTCTCTCAGAGCAACCCCTA 1439
 QY 1201 GTCTCAGGTGGAGAGCTGGCTCTCAGAGCTGTGCTTATCCAGAGATGAATAGAG 1260
 Db GTCTCAGGTGGAGAGCTGGCTCTCAGAGCTGTGCTTATCCAGAGATGAATAGAG 1499
 QY 1440 GTCTCAGGTGGAGAGCTGGCTCTCAGAGCTGTGCTTATCCAGAGATGAATAGAG 1499
 Db GTCTCAGGTGGAGAGCTGGCTCTCAGAGCTGTGCTTATCCAGAGATGAATAGAG 1499
 QY 1261 GGCAGGTGCTCTCCAGAGAGCAACAGTACATCTCTTCACTTGAAGCTGCAACAGGC 1320
 Db GGCAGGTGCTCTCCAGAGAGCAACAGTACATCTCTTCACTTGAAGCTGCAACAGGC 1559
 QY 1500 GGCAGGTGCTCTCCAGAGAGCAACAGTACATCTCTTCACTTGAAGCTGCAACAGGC 1559
 Db GGCAGGTGCTCTCCAGAGAGCAACAGTACATCTCTTCACTTGAAGCTGCAACAGGC 1559
 QY 1321 ATCCGCTGGAAGATCTCTTATATCAACAACAAGGAGCCCTTGTGAATCCCGAT 1380
 Db ATCCGCTGGAAGATCTCTTATATCAACAACAAGGAGCCCTTGTGAATCCCGAT 1619
 QY 1560 ATCCGCTGGAAGATCTCTTATATCAACAACAAGGAGCCCTTGTGAATCCCGAT 1619
 Db ATCCGCTGGAAGATCTCTTATATCAACAACAAGGAGCCCTTGTGAATCCCGAT 1619
 QY 1381 GGAATCTCTGCAATTAACAACCAACCAACCACTCAGAGCCCTGGAAGCCCATGTG 1440
 Db GGAATCTCTGCAATTAACAACCAACCAACCACTCAGAGCCCTGGAAGCCCATGTG 1679
 QY 1620 GGAATCTCTGCAATTAACAACCAACCAACCACTCAGAGCCCTGGAAGCCCATGTG 1679
 Db GGAATCTCTGCAATTAACAACCAACCAACCACTCAGAGCCCTGGAAGCCCATGTG 1679
 QY 1441 GGGCAGTCCCAACAGAGCCGCAAGAGAGAGCCCTCCGAGAGCCCAACAGAGGTC 1500
 Db GGGCAGTCCCAACAGAGCCGCAAGAGAGAGCCCTCCGAGAGCCCAACAGAGGTC 1739
 QY 1680 GGGCAGTCCCAACAGAGCCGCAAGAGAGAGCCCTCCGAGAGCCCAACAGAGGTC 1739
 Db GGGCAGTCCCAACAGAGCCGCAAGAGAGAGCCCTCCGAGAGCCCAACAGAGGTC 1739
 QY 1501 CAGCCACGAGAGCCCAAGATGAGAGCCCTCTGTCATCTCACTTGTCCAGGGGCTGAC 1560
 Db CAGCCACGAGAGCCCAAGATGAGAGCCCTCTGTCATCTCACTTGTCCAGGGGCTGAC 1799
 QY 1740 CAGCCACGAGAGCCCAAGATGAGAGCCCTCTGTCATCTCACTTGTCCAGGGGCTGAC 1799
 Db CAGCCACGAGAGCCCAAGATGAGAGCCCTCTGTCATCTCACTTGTCCAGGGGCTGAC 1799
 QY 1561 GCTTCTCCAGTCAAGTCAATCCGAGAGTCTTTTCCCTCCAGAGAGCTCTCACT 1620
 Db GCTTCTCCAGTCAAGTCAATCCGAGAGTCTTTTCCCTCCAGAGAGCTCTCACT 1859
 QY 1800 GCTTCTCCAGTCAAGTCAATCCGAGAGTCTTTTCCCTCCAGAGAGCTCTCACT 1859
 Db GCTTCTCCAGTCAAGTCAATCCGAGAGTCTTTTCCCTCCAGAGAGCTCTCACT 1859
 QY 1621 GTGTCTCAACGAGAGACTTTTCCATGAGAGATGATGAGCAACAGATTTGGCCGATG 1680
 Db GTGTCTCAACGAGAGACTTTTCCATGAGAGATGATGAGCAACAGATTTGGCCGATG 1919
 QY 1860 GTGTCTCAACGAGAGACTTTTCCATGAGAGATGATGAGCAACAGATTTGGCCGATG 1919
 Db GTGTCTCAACGAGAGACTTTTCCATGAGAGATGATGAGCAACAGATTTGGCCGATG 1919
 QY 1681 ACCCTGAGCCGAGATCTCTGGGGGAGAGCTCCTGAACCCCAATCGAGTCTGTATCCA 1740
 Db ACCCTGAGCCGAGATCTCTGGGGGAGAGCTCCTGAACCCCAATCGAGTCTGTATCCA 1979
 QY 1920 ACCCTGAGCCGAGATCTCTGGGGGAGAGCTCCTGAACCCCAATCGAGTCTGTATCCA 1979
 Db ACCCTGAGCCGAGATCTCTGGGGGAGAGCTCCTGAACCCCAATCGAGTCTGTATCCA 1979

QY	181	GGAAAAGGTAAC	TGA	CTCGGAG	CTGTCGTG	TGAGGAA	ATCTTCCG	CAGAC	CAAGGA	240						
Db	181	GGAAAAGGTAAC	TGA	CTCGGAG	CTGTCGTG	TGAGGAA	ATCTTCCG	CAGAC	CAAGGA	240						
QY	241	GGTGAATCTCTT	CAGGAT	CAGGAAT	CAATTCAT	TATTA	CAGCTTTCCAG	TTTTCAG	CTG	300						
Db	241	GGTGAATCTCTT	CAGGAT	CAGGAAT	CAATTCAT	TATTA	CAGCTTTCCAG	TTTTCAG	CTG	300						
QY	301	CAAGGAGGAT	TAATCTT	AGAAAGT	AGTACTCTG	AAGAGAA	AAAAAGAA	AGATTA	AAAC	360						
Db	301	CAAGGAGGAT	TAATCTT	AGAAAGT	AGTACTCTG	AAGAGAA	AAAAAGAA	AGATTA	AAAC	360						
QY	361	AAAGTA	TAACCTCTG	AAAAA	CCCAAGT	TCGAATG	TATCA	AAAAAGAT	TGACG	CAAGAA	420					
Db	361	AAAGTA	TAACCTCTG	AAAAA	CCCAAGT	TCGAATG	TATCA	AAAAAGAT	TGACG	CAAGAA	420					
QY	421	TACACG	ATTTCTAC	AGCAT	AGACTTA	CACAGAT	TTTCAT	TTAA	CACATTTAA	GAATAT	480					
Db	421	TACACG	ATTTCTAC	AGCAT	AGACTTA	CACAGAT	TTTCAT	TTAA	CACATTTAA	GAATAT	480					
QY	481	TCGAGG	GCAGAGT	ATAC	CTTTTG	AAAAATG	AGACG	AAAAATTA	TGATTTCA	TTCGTGAC	540					
Db	481	TCGAGG	GCAGAGT	ATAC	CTTTTG	AAAAATG	AGACG	AAAAATTA	TGATTTCA	TTCGTGAC	540					
QY	541	AACATAT	ATCATTTA	AAAAAGT	TTCCCTC	AGATG	TGTCAT	GTAT	TCAGAG	ATCCTGTGCAT	600					
Db	541	AACATAT	ATCATTTA	AAAAAGT	TTCCCTC	AGATG	TGTCAT	GTAT	TCAGAG	ATCCTGTGCAT	600					
QY	601	CGAGTGG	CAGCTTATTTT	TGGAT	TGGAT	TGATG	ATG	GCAT	CAAA	CAGGAAAAATCTG	TATC	660				
Db	601	CGAGTGG	CAGCTTATTTT	TGGAT	TGGAT	TGATG	ATG	GCAT	CAAA	CAGGAAAAATCTG	TATC	660				
QY	661	ATCAACA	AGAC	CCAGC	AGCACCAG	ATAC	CACAG	CCAAAG	TTTTTG	TGAACATTTA	AAAAAT	720				
Db	661	ATCAACA	AGAC	CCAGC	AGCACCAG	ATAC	CACAG	CCAAAG	TTTTTG	TGAACATTTA	AAAAAT	720				
QY	721	GAAAAAG	GTGAAG	AT	CCCAAG	AGCGG	TTTAT	CTTTGA	AG	CGAGAT	TA	780				
Db	721	GAAAAAG	GTGAAG	AT	CCCAAG	AGCGG	TTTAT	CTTTGA	AG	CGAGAT	TA	780				
QY	781	AAAGAG	CAAT	CA	AGT	CTG	TTG	CTCCAG	GA	AAAG	CC	840				
Db	781	AAAGAG	CAAT	CA	AGT	CTG	TTG	CTCCAG	GA	AAAG	CC	840				
QY	841	TTGGAGA	CAGT	TAACAT	TGCAAT	TGAGACT	TA	TAA	AAAAA	ACA	CAGCTCTT	TCGGGGCAAC	900			
Db	841	TTGGAGA	CAGT	TAACAT	TGCAAT	TGAGACT	TA	TAA	AAAAA	ACA	CAGCTCTT	TCGGGGCAAC	900			
QY	901	AGAGAT	GGCTCA	GGGAG	AACAT	CTTGG	AGT	CGAC	AGAC	AGCTCA	GA	AAAAATGA	ATGA	960		
Db	901	AGAGAT	GGCTCA	GGGAG	AACAT	CTTGG	AGT	CGAC	AGAC	AGCTCA	GA	AAAAATGA	ATGA	960		
QY	961	TGGTCTGA	CCAC	CAAAAG	GGCCTT	GGAG	CGAC	ACA	CTCCG	ACAT	GGTTCC	AAACCGGA	ATCTA	1020		
Db	961	TGGTCTGA	CCAC	CAAAAG	GGCCTT	GGAG	CGAC	ACA	CTCCG	ACAT	GGTTCC	AAACCGGA	ATCTA	1020		
QY	1021	AAAGCC	CGCAT	TGACCA	AGAC	GGCAG	AGTTT	TGG	GGG	GCAT	CA	CGTGTCT	GAC	CAAGGGGTAC	1080	
Db	1021	AAAGCC	CGCAT	TGACCA	AGAC	GGCAG	AGTTT	TGG	GGG	GCAT	CA	CGTGTCT	GAC	CAAGGGGTAC	1080	
QY	1081	AGCACTT	CCAGT	ACTAG	AGGAT	ACCG	GAAG	CTGT	CCAA	AG	CAAG	AGTTTCC	GA	CTTCCAGC	1140	
Db	1081	AGCACTT	CCAGT	ACTAG	AGGAT	ACCG	GAAG	CTGT	CCAA	AG	CAAG	AGTTTCC	GA	CTTCCAGC	1140	
QY	1141	AGTGA	AGGCTCT	CA	GGAT	TGCTGT	CCG	CA	CCCAT	TCAC	CTTCC	AGAG	CA	ACCCTTA	1200	
Db	1141	AGTGA	AGGCTCT	CA	GGAT	TGCTGT	CCG	CA	CCCAT	TCAC	CTTCC	AGAG	CA	ACCCTTA	1200	
QY	1201	GTCTCAG	GTGTG	GCAG	CTGT	CTTCA	CGCTGT	GTG	CTTAT	TCAG	AGAA	TG	GAAT	AGG	1260	
Db	1201	GTCTCAG	GTGTG	GCAG	CTGT	CTTCA	CGCTGT	GTG	CTTAT	TCAG	AGAA	TG	GAAT	AGG	1260	
QY	1261	GGCCAG	GTGTGCT	CCAG	CA	GCAC	CA	AGCTA	CA	TCTCTT	CCA	CTTGA	AG	CTGCA	CAAGC	1320

Db	1261	GGCCGAGTTGCTCCAGCAGCACCGCTACATCTCTTCCACTTGAACTGCAACAGGC	1320
QY	1321	ATCCCGCTGAAAGCATCTTCTTATATCCACACAGAGCCAGCCCTTGTGAATCCCGAT	1380
Db	1321	ATCCCGCTGGAAGCATCTTCTTATATCCACACAGAGCCAGCCCTTGTGAATCCCGAT	1380
QY	1381	GGAACTCTTGCATTTATACACCCACCACCACTCAGCAGCCCTGGGAAGCCGCATAGTG	1440
Db	1381	GGAACTCTTGCATTTATACACCCACCACCACTCAGCAGCCCTGGGAAGCCGCATAGTG	1440
QY	1441	GGGCGAGTCCCAACAGCAGCCCGCCACAGCAGCAGCCCTCCCGCAGCCCAACAGCAGTTC	1500
Db	1441	GGGCGAGTCCCAACAGCAGCCCGCCACAGCAGCAGCCCTCCCGCAGCCCAACAGCAGTTC	1500
QY	1501	CAGCCACCGCAGCCACAGATGGCAGAGCCCTTGATCACTCAGTCTGTCCAGGGGCTGCAG	1560
Db	1501	CAGCCACCGCAGCCACAGATGGCAGAGCCCTTGATCACTCAGTCTGTCCAGGGGCTGCAG	1560
QY	1561	GCTTCTCTCCCATCAGTGCATATATCCGGCAGTCTCTTTTCTCTCCCAAGACCTCTCACT	1620
Db	1561	GCTTCTCTCCCATCAGTGCATATATCCGGCAGTCTCTTTTCTCTCCCAAGACCTCTCACT	1620
QY	1621	GTGTCTCCAAAGCAGCAGCTTTCCATGAGATGATGTGGCAACAGTGTGGCCAGATG	1680
Db	1621	GTGTCTCCAAAGCAGCAGCTTTCCATGAGATGATGTGGCAACAGTGTGGCCAGATG	1680
QY	1681	ACCCTGAGCCGGCAGTCTCTGGGGGAGACTCTGTGAACCCCAATCAGTCTGTCTACCCA	1740
Db	1681	ACCCTGAGCCGGCAGTCTCTGGGGGAGACTCTGTGAACCCCAATCAGTCTGTCTACCCA	1740
QY	1741	TCCTCTCTTATGCAACAGCCGGCCACGACGCCAGCTATGTATGCTCTTACAGGCCAG	1800
Db	1741	TCCTCTCTTATGCAACAGCCGGCCACGACGCCAGCTATGTATGCTCTTACAGGCCAG	1800
QY	1801	CAGCTTCCCTAAGAGGGAATTCCTCAGGCTGTGGCCCTCCATCTCCCGCAGAGGCTCTCCAG	1860
Db	1801	CAGCTTCCCTAAGAGGGAATTCCTCAGGCTGTGGCCCTCCATCTCCCGCAGAGGCTCTCCAG	1860
QY	1861	CCCCCTCCCTCACCACAGGGAATCTGTGCAACAGCCTTCGCGCTGCACAGATGCTGTATAT	1920
Db	1861	CCCCCTCCCTCACCACAGGGAATCTGTGCAACAGCCTTCGCGCTGCACAGATGCTGTATAT	1920
QY	1921	TATTATCCCATGTGTGTCAGTACCTTACCTCACAACGCAACAGTACCGGCCATGCCCCG	1980
Db	1921	TATTATCCCATGTGTGTCAGTACCTTACCTCACAACGCAACAGTACCGGCCATGCCCCG	1980
QY	1981	GTTTACGTACAAACGCTCAGAAGAGTAAACAGAGTGCACAGGCGACACAGGAAGAGGTTAC	2040
Db	1981	GTTTACGTACAAACGCTCAGAAGAGTAAACAGAGTGCACAGGCGACACAGGAAGAGGTTAC	2040
QY	2041	CAGCCAGCTTGTCTGTGTCACACAGGGAATTCCAAGGCTCTAATGAGGTGACAGCGCACT	2100
Db	2041	CAGCCAGCTTGTCTGTGTCACACAGGGAATTCCAAGGCTCTAATGAGGTGACAGCGCACT	2100
QY	2101	CAGAGTCAAGAACTGTATTAATTAACCAACAGGAAGTCCGGTGCAAAGCGTATGATTTC	2160
Db	2101	CAGAGTCAAGAACTGTATTAATTAACCAACAGGAAGTCCGGTGCAAAGCGTATGATTTC	2160
QY	2161	TACCCAAACATGTCTTCTTATTCAGGTGCCAATGACCCAGGGTTCTTAAGAATGCCCCAG	2220
Db	2161	TACCCAAACATGTCTTCTTATTCAGGTGCCAATGACCCAGGGTTCTTAAGAATGCCCCAG	2220
QY	2221	CAGTCATACCAACAGCCATCATGTCTTACCTTAACAGGCAAGTCTAAGGTCACTCCAGCC	2280
Db	2221	CAGTCATACCAACAGCCATCATGTCTTACCTTAACAGGCAAGTCTAAGGTCACTCCAGCC	2280
QY	2281	ACTGGAAATGCTGTTTACTGTATGTCAACCGGCCACCCCTCAGAACCACTTAGGCTG	2340
Db	2281	ACTGGAAATGCTGTTTACTGTATGTCAACCGGCCACCCCTCAGAACCACTTAGGCTG	2340
QY	2341	ATTGGCCCAACTGCCCCCTCCAGCATGTCTCCAGTGAATGTACGTAGCTGCAGAACAAAC	2400
Db	2341	ATTGGCCCAACTGCCCCCTCCAGCATGTCTCCAGTGAATGTACGTAGCTGCAGAACAAAC	2400

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: July 5, 2006, 09:50:18 ; Search time 32 Seconds

(without alignments)
681.702 Million cell updates/sec

Title: US-10-509-950-1

Perfect score: 4235
Sequence: 1 MSEQGDINQAIABEGTQOE.....SASCRITNCASMSNAGQVKF 813

Scoring table:

BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 112942 seqs, 26832045 residues

Total number of hits satisfying chosen parameters: 112942

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database:

1: /BMC_Celerra_SIDS3/prodata/1/pubppa/US09_NEW_PUB.ppe:*
2: /BMC_Celerra_SIDS3/prodata/1/pubppa/US06_NEW_PUB.ppe:*
3: /BMC_Celerra_SIDS3/prodata/1/pubppa/US07_NEW_PUB.ppe:*
4: /BMC_Celerra_SIDS3/prodata/1/pubppa/US08_NEW_PUB.ppe:*
5: /BMC_Celerra_SIDS3/prodata/1/pubppa/FCI_NEW_PUB.ppe:*
6: /BMC_Celerra_SIDS3/prodata/1/pubppa/US10_NEW_PUB.ppe:*
7: /BMC_Celerra_SIDS3/prodata/1/pubppa/US11_NEW_PUB.ppe:*
8: /BMC_Celerra_SIDS3/prodata/1/pubppa/US60_NEW_PUB.ppe:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	621	14.7	637	US-11-289-102-335	Sequence 335, App
2	318.5	7.5	875	US-10-449-902-41298	Sequence 41298, A
3	232.5	5.5	1034	US-10-449-902-41078	Sequence 41078, A
4	229	5.4	2245	US-11-248-956-3	Sequence 3, Appl
5	229	5.4	2414	US-11-283-329-176	Sequence 176, App
6	226.5	5.3	595	US-10-449-902-38085	Sequence 38085, A
7	222	5.2	594	US-10-449-902-52090	Sequence 52090, A
8	222	5.2	1189	US-10-449-902-41223	Sequence 41223, A
9	221	5.2	443	US-10-953-349-13468	Sequence 13468, A
10	221	5.2	451	US-10-953-349-13467	Sequence 13467, A
11	221	5.2	456	US-10-953-349-13467	Sequence 13467, A
12	215	5.1	693	US-11-293-697-3849	Sequence 3849, App
13	212	5.0	554	US-10-953-349-23196	Sequence 23196, A
14	211.5	5.0	1073	US-10-449-902-41088	Sequence 41088, A
15	210	5.0	1252	US-10-449-902-55024	Sequence 55024, A
16	205.5	4.9	802	US-10-527-572-5	Sequence 5, Appl
17	204	4.8	389	US-10-449-902-33007	Sequence 33007, A
18	204	4.8	1464	US-11-283-329-152	Sequence 152, App
19	203.5	4.8	295	US-10-953-349-4000	Sequence 4000, App
20	203.5	4.8	402	US-10-953-349-3999	Sequence 3999, App
21	203	4.8	418	US-10-953-349-23197	Sequence 23197, A
22	203	4.8	2517	US-11-283-329-204	Sequence 204, App
23	202.5	4.8	536	US-10-449-902-40167	Sequence 40167, A
24	201.5	4.8	683	US-10-449-902-28582	Sequence 28582, A
25	200.5	4.7	804	US-10-486-020-11	Sequence 11, Appl

26	200.5	4.7	1145	US-10-997-086-5	Sequence 5, Appl
27	200	4.7	963	US-10-142-275B-4	Sequence 4, Appl
28	200	4.7	1595	US-10-486-020-18	Sequence 18, Appl
29	199.5	4.7	905	US-10-449-902-41192	Sequence 41192, A
30	199	4.7	1090	US-10-530-450-10	Sequence 10, Appl
31	198.5	4.7	389	US-10-953-349-12088	Sequence 12088, A
32	198	4.7	461	US-10-449-902-51694	Sequence 51694, A
33	197.5	4.7	268	US-10-953-349-4001	Sequence 4001, App
34	197.5	4.7	966	US-10-449-902-41240	Sequence 41240, A
35	197	4.7	1042	US-10-449-902-52688	Sequence 52688, A
36	196.5	4.6	3924	US-11-165-586-20	Sequence 20, Appl
37	195.5	4.6	546	US-10-449-902-28733	Sequence 28733, A
38	195	4.6	346	US-10-953-349-12089	Sequence 12089, A
39	194.5	4.6	634	US-11-293-697-3340	Sequence 3340, App
40	194.5	4.6	738	US-10-449-902-51458	Sequence 51458, A
41	194.5	4.6	964	US-10-142-275B-2	Sequence 2, Appl
42	194	4.6	1078	US-10-449-902-54534	Sequence 54534, A
43	193	4.6	2124	US-11-283-329-160	Sequence 160, App
44	192.5	4.5	375	US-10-953-349-23198	Sequence 23198, A
45	192	4.5	343	US-10-953-349-21437	Sequence 21437, A

ALIGNMENTS

RESULT 1	US-11-289-102-335
Sequence 335, Application US/11289102	
Publication No. US20060121511A1	
GENERAL INFORMATION:	
APPLICANT: Lee, Hyelim	
APPLICANT: Shaw, Peter M.	
TITLE OF INVENTION: BLOWERS AND METHODS FOR DETERMINING SENSITIVITY TO MICROBIAL STABILIZING AGENTS	
FILE REFERENCE: 10338 NP	
CURRENT APPLICATION NUMBER: US/11/289, 102	
CURRENT FILING DATE: 2005-11-29	
PRIOR APPLICATION NUMBER: US 60/631, 993	
PRIOR FILING DATE: 2004-11-30	
NUMBER OF SEQ ID NOS: 395	
SOFTWARE: PatentIn version 3.3	
SEQ ID NO 335	
LENGTH: 637	
TYPE: PRT	
ORGANISM: Homo sapiens	
US-11-289-102-335	
Query Match	14.7%; Score 621; DB 7; Length 637;
Best Local Similarity	35.3%; Pred. No. 4.5e-22;
Matches 157; Conservative 60; Mismatches 96; Indels 132; Gaps 16;	
QY	340 LKPMYKASFGITVTRDSTSTSTSTGKLSKASBSSSSAGSGSLRTHPPLOSTP 399
	1 MRPPVTKASFSFGISITRBDISI-----GSKGSGAGISR----- 36
DB	400 LVGVNAGSGCVPYFENGIGGVAPSTSVILLPLAATIPGSLINHTGQPFVNP 459
QY	37 -----FGMAL----- 41
DB	460 DGTPIATNPPTSOQLPLASAVGSGQQPPOQPPPOQVPPPOPMAGPLVYQSVGL 519
QY	42 -GAPEVGNQVTSQSVGLPLPCTAQOQQOQO-----QQQLPALPPTPOQPPPLNHHMI--- 93
DB	520 QASQSVQYAVSPFPFHLLPVSTQHPFMRDVIATQFGQNTLSRSGSGETPEPSPGVY 579
QY	94 ---SQA-----DLSNPFQWLSLRQSTELAD-PSAALF 124
DB	580 PSSIMPPOPAQPSVIASTGOQLPTGFGS-GSGPPISQVLYQPPSPQGFVQPPPAOMP 638
QY	125 QTPILSHPPQOTSFIMASTGOPLPTSVYTSSSHAPFTQVUL-----PFGVWQ--PPOIQ 178
DB	639 VVYVPSGQVPTSTTQOYRPM-A-PVQYNAQBSQQQMPQAAQAGVQVPLSGQO--GFGGLIG 695

179 VSYPPGQVFN--NQYRPLSHFVAXSPRGQQLQPSQQLQPMMPQQAQMGITG 237
636 VQGPQSQNVINNOQ--GTPVQSVWVSYPTNYSYQVMTQSGQLPQGSYQOPIMLP-N 751
238 VQO--PQNGGLSSQSSMGGQMGQLVQYTPSPSYQVPGVSDSQNVVQPPQPMPLVPVS 296
752 QAGGSLPATGMFYCANTPTPTPON 776
297 QSVGGPLPAGVFPVYYSMTIPPAQON 321

RESULT 2

US-10-449-902-41298
Sequence 41298, Application US/10449902
Publication No. US20060123505A1
GENERAL INFORMATION:
APPLICANT: National Institute of Agricultural Sciences.
APPLICANT: Bio-oriented Technology Research Advancement Institution.
APPLICANT: The Institute of Physical and Chemical Research.
TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
FILE REFERENCE: MOA-A0205Y1-US
CURRENT APPLICATION NUMBER: US/10/449,902
CURRENT FILING DATE: 2003-05-29
PRIOR APPLICATION NUMBER: JP 2002-203269
PRIOR FILING DATE: 2002-05-30
PRIOR APPLICATION NUMBER: JP 2002-383870
PRIOR FILING DATE: 2002-12-11
NUMBER OF SEQ ID NOS: 56791
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 41298
LENGTH: 875
TYPE: PRT
ORGANISM: Oryza sativa
US-10-449-902-41298

Query Match 7.5%; Score 318.5; DB 6; Length 875;
Best Local Similarity 23.1%; Pred. No. 6,2e-08; Indels 319; Gaps 47;
Matches 207; Conservative 95; Mismatches 276;

8 NOAIEBGEQTEATPENGIVKESLDEEKELELQRLLEAQNORRKSAGK----- 62
109 SRTVLTSGGLQPTSTDEGTLLS-----SSDDGKPSLD 144
63 GKLTFS--LAVCESSAPPGESLQDOESHILQLSFSSLOEEDKSRKDSEREKDK 119
145 GKSVAAGTTFALDEKESLRP-----DSSASV-----QASAEED--DL 179
120 NKDKTSEKPKIRMLSKDCSQEYTDSTGIDLHEF----- 152
180 CYPGTGVPITSRMGSBDGVAFRD--QLHEIHAMEPRRQAGPAGANGALPQGLVYP 235
153 -----LINTLKNNSRDMILLKMEQEIIFDIADNN 182
226 PEGPVGSIIPRSRQALATGGAAYAPDKLEALE-NPRDIWVLEQDVIVDKAK 294
183 NHYKFPQMSYORMLVHRVAAVFGLDHNVDTGKSVIINKTSSTRIP-----QRF 235
295 ESSLMLPQONSFRMLAHKIADYVMLGSHVSDSNSAVRLFRTPHCRIPQPLTGVTPTA 354
226 EHLKQKEEEOKRILKRDNSIDKEDQSVCSQESLFVENSRLLEDNSINENYKKG 295
335 ASTPPSGPQIK--LKR--GID--PGPATANGSNV--NSKTNSDDDSGD--KKRP 401
296 LFRGND-----GSGRTSGROSSSENEKMSDHPQAMSTDSSENRLKPM 344
402 VTREERARVAAARLRIMGSAKPTGSETSKED-----DERSSSAAGKSK 447
345 TKTAAPGGLTVITRDSTSTSTGKLSKAGSESSSSAGSSSLRTHPPLQSTLVAGV 404
448 KK-----QRNDSDDGEFARSAYSSYFAPSSSAGYT-----SQVAYONQGTVA 492

405 AAGSPGVPEENGIGGVAPSSSTSYLLPLEATGIPGSLILNPH-TGQFVNPDGTP 463
493 AOFSSS-PT-----GNDDMAASPY-----GVQONQ--NSHSGSYTSGESG 533
464 AINPP-----TSQQLRS--AMV-----GQS--QQQPPQQ 491
534 SMYQVPGQYDLSKQFNQMSFQNSMADQQQPVYSFYAQQVYAGSGYSSAMQHQHQ 593
492 PSPPQQQVQ--PPQPMAG-----PLVT-QSYQGLQASQSQYQYAVSP 534
594 QSOHQOQOQCGYNATQPPSTFGSPSPNSDRNQMPVYTSQOQOQOQLHQOQOQNPFGQLP 653
535 PQ-----HLTPVS-PTQHPFMRDV-----ATQFGQMTLSRQSGETPE-PPS 575
654 SONTGRVPSALEHPLPWSKSPHNPQSTVPPGPIGNSNSSFPPQAPATGNSNNGHQ 713
576 GPVYPSLMPQPAQOPSYVIATGQOL--PTGFGSGGP--ISQOVLQPPSPQGFVQ 631
714 G-VPSLERQMSQ-----SQGSAPGSPHHGMHAGGAPMPSQPLTHLPQPV-FPRQ 764
632 PPP-----AQPVYVYPSGQVPTSTTQO-----YRMAPVQ- 662
765 PSPVVALPPEKISGHQWPLEH--ARQSPSNVHQMSPOSSSIKMGAPASLEKPPSPVP 822
663 YNAOR--SQOMPQAQAGYQPVYLSGQGFQGLIGVQOPQSQNVINNOGTVPYOS 716
823 FDASRFTSQTPSPFNMAAARLPGGGLSGF-----GPMQPLNSSGMANG-NQSPRS 874

RESULT 3

US-10-449-902-41078
Sequence 41078, Application US/10449902
Publication No. US20060123505A1
GENERAL INFORMATION:
APPLICANT: National Institute of Agricultural Sciences.
APPLICANT: Bio-oriented Technology Research Advancement Institution.
APPLICANT: The Institute of Physical and Chemical Research.
TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
FILE REFERENCE: MOA-A0205Y1-US
CURRENT APPLICATION NUMBER: US/10/449,902
CURRENT FILING DATE: 2003-05-29
PRIOR APPLICATION NUMBER: JP 2002-203269
PRIOR FILING DATE: 2002-05-30
PRIOR APPLICATION NUMBER: JP 2002-383870
PRIOR FILING DATE: 2002-12-11
NUMBER OF SEQ ID NOS: 56791
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 41078
LENGTH: 1034
TYPE: PRT
ORGANISM: Oryza sativa
US-10-449-902-41078

Query Match 5.5%; Score 232.5; DB 6; Length 1034;
Best Local Similarity 20.1%; Pred. No. 0.0007;
Matches 193; Conservative 99; Mismatches 357; Indels 309; Gaps 43;

2 SEQDLNOAIA--EEGTEQETATPENGIVKESLDEEKELELQRLLEAQNORRKS 58
158 TEGELGVALASAOQIPADAKDAQPOLGALKSSLLINAKLESIOAEVVRVODAKEMR 217
59 GA-----GKGLTR-----SLAVCESSAR-----PG 81
218 EAREKEAFRTTISKAFDPLGSGRALQKMSDDPLALDDHSSLRPMASPSDAKASGA 277
82 ESLDOES-----IHLQLSFSSLOEE--DKSRKDSEREKDK----- 119
278 GLEDSNKNKPMQDPAANEARFGIKSNYEENLYTKLDRTGKDFDRERADRIRAREING 337
120 ---NKDKTSEKPKIRMLSKDCSQEYTDSTGIDLHEFLINTLKNNSRDMILLKMEQEI 176
338 QATNPPHMAEB-----RGHADDSGVN-----BEDYKG 364

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: July 5, 2006, 09:49:08 ; Search time 31 Seconds

(without alignments)
2295.561 Million cell updates/sec

Title: US-10-509-950-1

Sequence: 1 MSEGQDLMQAIABEGSTGEF.....SASCRTNCASMSNAGVQVRF 813

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database:

Issued Patents AA:*

- 1: /EMC_Celerra_SIDS3/ptodata/2/1aa/5-COMB.pep:*
- 2: /EMC_Celerra_SIDS3/ptodata/2/1aa/6-COMB.pep:*
- 3: /EMC_Celerra_SIDS3/ptodata/2/1aa/7-COMB.pep:*
- 4: /EMC_Celerra_SIDS3/ptodata/2/1aa/H-COMB.pep:*
- 5: /EMC_Celerra_SIDS3/ptodata/2/1aa/PCUTS-COMB.pep:*
- 6: /EMC_Celerra_SIDS3/ptodata/2/1aa/RE-COMB.pep:*
- 7: /EMC_Celerra_SIDS3/ptodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1510	35.7	971	2	US-09-538-092-1332 Sequence 1332, Ap
2	473	11.2	107	2	US-10-104-047-3736 Sequence 3736, Ap
3	293	6.9	1444	2	US-09-949-016-9652 Sequence 9652, Ap
4	293	6.9	2004	2	US-09-538-092-1371 Sequence 1371, Ap
5	293	6.9	2004	2	US-09-949-016-6756 Sequence 6756, Ap
6	250	5.9	1185	2	US-09-041-886-23 Sequence 23, Appl
7	250	5.9	1185	2	US-09-538-092-1209 Sequence 1209, Ap
8	246.5	5.9	1184	2	US-09-266-225D-18 Sequence 18, Appl
9	243.5	5.7	369	1	US-08-991-307-2 Sequence 2, Appl
10	241	5.7	131	2	US-09-270-767-44309 Sequence 44309, A
11	233.5	5.5	2185	2	US-09-854-856-36 Sequence 36, Appl
12	233.5	5.5	2185	2	US-10-010-720-36 Sequence 36, Appl
13	233.5	5.5	2245	2	US-09-854-856-4 Sequence 4, Appl
14	233.5	5.5	2245	2	US-10-010-720-4 Sequence 4, Appl
15	233.5	5.5	2322	2	US-09-854-856-34 Sequence 34, Appl
16	233.5	5.5	2322	2	US-10-010-720-34 Sequence 34, Appl
17	233.5	5.5	2382	2	US-09-854-856-2 Sequence 2, Appl
18	233.5	5.5	2382	2	US-10-010-720-2 Sequence 2, Appl
19	232	5.5	2442	2	US-09-514-247A-10 Sequence 10, Appl
20	232	5.5	2442	2	US-09-538-092-1370 Sequence 1370, Ap
21	229	5.4	2414	1	US-08-227-536-2 Sequence 2, Appl
22	229	5.4	2414	1	US-09-538-092-1289 Sequence 1289, Ap
23	229	5.4	2414	5	PCR-US95-04682-2 Sequence 2, Appl
24	227	5.4	1345	3	US-10-433-794-17 Sequence 17, Appl
25	227	5.4	579	2	US-09-668-119-3 Sequence 3, Appl
26	224.5	5.3	922	2	US-09-854-856-38 Sequence 38, Appl

27	224.5	5.3	922	2	US-10-010-720-38 Sequence 38, Appl
28	224.5	5.3	982	2	US-09-854-856-6 Sequence 6, Appl
29	224.5	5.3	982	2	US-10-010-720-6 Sequence 6, Appl
30	222	5.2	2441	1	US-08-194-468-2 Sequence 2, Appl
31	222	5.2	2441	2	US-08-261-739-2 Sequence 2, Appl
32	222	5.2	2441	2	US-09-514-247A-8 Sequence 8, Appl
33	222	5.2	2441	2	US-09-866-316-2 Sequence 2, Appl
34	221.5	5.2	802	2	US-09-823-240A-2 Sequence 2, Appl
35	221.5	5.2	2079	2	US-09-949-016-8301 Sequence 8301, Ap
36	219.5	5.2	1253	2	US-09-489-039A-12097 Sequence 12097, A
37	219	5.2	1326	2	US-09-688-188B-15 Sequence 15, Appl
38	219	5.2	1326	2	US-09-291-417D-15 Sequence 15, Appl
39	215	5.1	1274	2	US-09-095-443-2 Sequence 2, Appl
40	215	5.1	1636	3	US-10-433-794-1 Sequence 1, Appl
41	213	5.0	1179	2	US-09-949-016-7088 Sequence 7088, Ap
42	212	5.0	915	2	US-10-029-180-74 Sequence 74, Appl
43	211.5	5.0	2142	2	US-09-538-092-1142 Sequence 1142, Ap
44	211.5	5.0	2142	2	US-09-949-002-371 Sequence 371, Appl
45	211.5	5.0	2453	2	US-09-632-033B-4 Sequence 4, Appl

ALIGNMENTS

```
RESULT 1
US-09-538-092-1332
Sequence 1332, Application US/09538092
Patent No. 6753314
GENERAL INFORMATION:
APPLICANT: Glot, Lotic
APPLICANT: Mansfield, Traci A.
TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
FILE REFERENCE: 15966-542
CURRENT APPLICATION NUMBER: US/09/538, 092
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: 60/127,352
PRIOR FILING DATE: 1999-04-01
PRIOR APPLICATION NUMBER: 60/178,965
PRIOR FILING DATE: 2000-02-01
NUMBER OF SEQ ID NOS: 1387
SOFTWARE: CurafastSeqFormatter Version 0.9
SEQ ID NO 1332
LENGTH: 971
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (0)...(0)
OTHER INFORMATION: Polypeptide Accession Number Q15032
US-09-538-092-1332
Query Match 35.7%; Score 1510; DB 2; Length 971;
Beet Local Similarity 40.6%; Pred. No. 1.5e-99;
Matches 332; Conservative 124; Mismatches 135; Indels 226; Gaps 23;
18 EOEETAPNGVYSSLSBEKLELQRLAONQRRKSKSGAGKGLTRSLAVEESS 76
10 KDETIT-----MKDLAEVKTQTRVENLTKSEN-----GKLVKQNEHTEN 52
77 ARPGESELODESIMLOUSFSFSLQEBDKSRKDSEREKDKDKTKSEKIRMLSKD 136
53 -----NIDLEKIQIQT--QSFPEKKEKPSDKBAEKEKASDK-----LRL-KQLSRD 96
137 CSQETDSTGIDHFLNTLKNKSRDMILKMEQETIDFLADNNNHKKRPFQSSYOR 196
97 SSQETDSTGIDHFLNTLKNKSRDMILKMEQETIDFLADNNNHKKRPFQSSYOR 156
197 MLVHRVAAYFGLDHVDQTKGKSVIINKTSSTRIPQRCFCEHLKDKGSESRFTIKRN 256
157 MLVHRVAAYFGLDHVDQTKGKSVIINKTSSTRIPQRCFCEHLKDKGSESRFTIKRN 216
257 SSIDKEDNQ-----SVCSQESLFLVENSRLLED 284
```

Db	217	SSFDKDDNDQVRRRLKLDKDRSKSLSEEEBEYQRRARDIIFQSDLSGSENYIID--KRLQDE	274
Qy	265	NICNETYKRRQLFRGNRDGSGRTSGSGSSSENELKWSHDQRWASSTDDSSNRNLKPM	344
Db	275	D-ASSTOQRRQIFRVRVKDASGRSTNSHGSTENELKYSE-PRWSSSTDSSSLRLKPAV	332
Qy	345	TTTASFGGITVLTFRGDSTSTRSTGKLSKAGSSSSSSAGSSSLSTTHPPLOSTPLVSGV	404
Db	333	TKASSPFGISVLTLRGSSSSSKSIGLSK----	361
Qy	405	AAGSPGCVYPENIGIGQVAPASTSYILLPLEATGIPGSIILNPHTGQPFVNPDGTPA	464
Db	362	-----TQPFINPDGSEV	374
Qy	465	IYNPTSGQPLRSAMVGSQOQPPQOQSPRPQOQVQPPQPMAGBLVTQSVGLQASSQ	524
Db	375	VYNPTMQQPVASQVVG-----PQPPPLPAPPQ-----	403
Qy	525	SVQYVAVSPFPQHLPLVSPTHQPMRDVATQCGMTLSRQSSGRTPEPPSGVYPPSSLM	584
Db	404	-----PAAN-----H-----FSQDNLGSGPSHSLRLARPADSDSDPA--MFQSTV	446
Qy	585	PPAQAQPSYVIAS-----GQQLPTGGSGSGSPISQOVLQPPPS	624
Db	447	LQSPQSGIMTAAPRRPPRRPPRRPPRLPRQGPVYTAGVPSAGHVSQPVLG----	502
Qy	625	PQGFVQPPPAQMPVYVYYPSSQYPTSTTQQRPMADVQVYNAORSQMPQAQOAGYQVPL	684
Db	503	QGGYVQOQESP--QMPACYCAPHY-HSSQPYRRPVSVHNHSLNQLPLPQAQGTGYQVIP	560
Qy	685	SGQGFQGLIGVQOPQOSQNVINNOQ-----GTFVQSVMSVYPMSSVQVMTQSGSLPQO	741
Db	561	NOQNVYQIGVQO--POSQSLVSGQPNLSGNQIQGVVITYSVPTVYQVSLPOSQGIPIHQ	619
Qy	742	SYQDPIMLPNQAGQSLPATGMVYCNVTPPTPNNL	778
Db	620	TYQGFVMPFNQSNQSMPTTGMPVYVSVIPEQGNLL	656
RESULT 2			
US-10-104-047-3736			
Sequence 3736, Application US/10104047			
Patent No. 6943241			
GENERAL INFORMATION:			
APPLICANT: HELIX RESEARCH INSTITUTE			
TITLE OF INVENTION: No. 6943241el full length cDNA			
FILE REFERENCE: H1-A0105			
CURRENT APPLICATION NUMBER: US/10/104, 047			
CURRENT FILING DATE: 2002-03-25			
PRIOR APPLICATION NUMBER:			
PRIOR FILING DATE:			
NUMBER OF SEQ ID NOS: 4096			
SOFTWARE: PatentIn Ver. 2.1			
SEQ ID NO 3736			
LENGTH: 107			
TYPE: prt			
ORGANISM: Homo sapiens			
US-10-104-047-3736			

```

Query Match          11.2% ; Score 473 ; DB 2 ; Length 107 ;
Best Local Similarity 86.5% ; Pred. No. 1.5e-26 ;
Matches 90 ; Conservative 2 ; Mismatches 4 ; Indels 8 ; Gaps 1

Qy 718 MVSYPYMSY-----QVPTQSGGIPQSGYOQPIPLPNOAGGSLPATGMPYCNV 769
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 4 IVQKRIVSRYLACVFLQVPTQSGGIPQSGYOQPIPLPNOAGGSLPATGMPYCNV 63

Qy 770 TTPPTONNLRLIGPHCPSSITVPVMSASGRTNCAASNMGQVKF 813
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 64 TTPPTONNLRLIGPHCPSSITVPVMSASGRTNCAASNMGQVKF 107

```

```

US-09-949-016-9652
/ Sequence 9652, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: C001307
/ CURRENT APPLICATION NUMBER: US/09/949, 016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 9652
/ LENGTH: 1444
/ TYPE: PRT
/ ORGANISM: Human
US-09-949-016-9652

```

[illegible]

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: July 5, 2006, 09:47:58 ; Search time 309 Seconds

(without alignments)
2433.779 Million cell updates/sec

Title: US-10-509-950-1

Perfect score: 4235
Sequence: 1 MSBQGLNQAIAEAGTEOE.....SASCRTNCASMSNAGQVXF 813

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : uniprot/2.2*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4111	97.1	793	2 Q4G0V4 HUMAN	Q4G0V4 homo sapien
2	3997.5	94.4	812	2 Q86VJ1 HUMAN	Q86VJ1 homo sapien
3	3641	86.0	809	2 Q7TS83 MOUSE	Q7TS83 mus musculu
4	3533	83.4	789	2 Q8C038 MOUSE	Q8C038 mus musculu
5	3433.5	81.1	771	2 Q8C018 MOUSE	Q8C018 mus musculu
6	3421	80.8	807	2 Q91Y59 MOUSE	Q91Y59 mus musculu
7	3414.5	80.6	651	2 Q9UF93 HUMAN	Q9UF93 homo sapien
8	2450	57.9	601	2 Q3UUF8 MOUSE	Q3UUF8 mus musculu
9	1831.5	43.2	784	2 Q53S01 HUMAN	Q53S01 homo sapien
10	1831.5	43.2	1099	1 R3HD1 HUMAN	015032 homo sapien
11	1827.5	43.2	1061	2 Q5R7K5 PONY	Q5R7K5 pongo pygma
12	1789	42.2	335	2 Q49A56 HUMAN	Q49A56 homo sapien
13	1570	37.1	1044	1 R3HD2 MOUSE	Q80T46 mus musculu
14	1493	35.3	962	1 R3HD2 HUMAN	Q9Y2K5 mus musculu
15	1463	34.5	1009	2 Q50516 XENLA	Q50516 xenopus lae
16	1316.5	31.1	989	2 Q4SS78 TETNG	Q4SS78 tetradon n
17	1298	30.6	823	2 Q5R8U0 PONY	Q5R8U0 pongo pygma
18	1296.5	30.6	549	2 Q80ZB9 MOUSE	Q80ZB9 mus musculu
19	1231	29.1	372	2 Q8C0Y6 MOUSE	Q8C0Y6 mus musculu
20	1109.5	26.2	466	2 Q6NRG3 XENLA	Q6NRG3 xenopus lae
21	1075.5	25.4	1066	2 Q4SE33 TETNG	Q4SE33 tetradon n
22	1036.5	24.5	248	2 Q8BMM1 MOUSE	Q8BMM1 mus musculu
23	1030.5	24.3	228	2 Q8BUD1 MOUSE	Q8BUD1 mus musculu
24	976.5	23.1	455	2 Q5PPF9 RAT	Q5PPF9 rattus norv
25	942.5	22.3	352	2 Q4R3F1 MACRA	Q4R3F1 macaca fasc
26	741	17.5	347	2 Q4RNM3 TETNG	Q4RNM3 tetradon n
27	736.5	17.4	681	2 Q3ZCT5 HUMAN	Q3ZCT5 homo sapien
28	679	16.0	657	2 Q5XKA9 MOUSE	Q5XKA9 mus musculu
29	672	15.9	656	2 Q2MTT9 HUMAN	Q2MTT9 homo sapien
30	652	15.4	445	2 Q95KJ7 MACRA	Q95KJ7 macaca fasc
31	617	14.6	239	2 Q3URW7 MOUSE	Q3URW7 mus musculu

ALIGNMENTS

RESULT 1	Q4G0V4 HUMAN	PRELIMINARY; PRT; 793 AA.	
AC	Q4G0V4		
DT	30-AUG-2005, integrated into UniProtKB/TREMBL.		
DT	07-FEB-2006, entry version 1.		
DE	ARPP-21 protein.		
GN	Name=ARPP-21;		
OS	Homo sapiens (human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae;		
OC	Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RC	TTSUB=Brain;		
RX	MDLLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.2426038999;		
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,		
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,		
RA	Altschul S.F., Zeeberg B., Buelow K.H., Schaefer C.F., Bhat N.K.,		
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,		
RA	Datchenko U., Marusina K., Farmer A.A., Rubin G.W., Hong L.,		
RA	Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,		
RA	Brownstein M.J., Usdin T.B., Tohilyuki S., Carninci P., Prange C.,		
RA	Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,		
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,		
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulys S.W.,		
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,		
RA	Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,		
RA	Willing M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,		
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,		
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,		
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,		
RA	Schmarch A., Schein J.E., Jones S.J.M., Marra W.A.,		
RT	"Generation and initial analysis of more than 15,000 full-length human		
RL	and mouse cDNA sequences."		
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).		
RP	NUCLEOTIDE SEQUENCE.		
RC	TTSUB=Brain;		
RG	NH MGC Project;		
RL	Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.		
CC	Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms		
CC	distributed under the Creative Commons Attribution-NonDerivative license		
CC	EMBL, BC036399, AA36399.1, -, mRNA.		
DR	SMR, Q4G0V4, 147-257.		
DR	GO, GO:0003676, F:uncle acid binding, IEA.		
DR	InterPro, IPR001374, R3H_ss.bd.		
DR	Pfam, PF01424, R3H, 1.		
DR	SMART, SM00393, R3H, 1.		
DR	PROSITE, PS51061, R3H, 1.		

32	430.5	10.2	108	2 Q6NYC3 HUMAN	Q6NYC3 homo sapien
33	430	10.2	89	1 ARP21 HUMAN	Q6B10 homo sapien
34	430	10.2	89	1 ARP21 PONY	Q5R6X9 pongo pygma
35	430	10.2	89	2 Q49AK3 HUMAN	Q49AK3 homo sapien
36	430	10.2	89	2 Q4R4W2 MACRA	Q4R4W2 macaca fasc
37	429	10.1	1738	2 Q2M075 DROPS	Q2M075 drosophila
38	414	9.8	1823	2 Q5BIH9 DROME	Q5BIH9 drosophila
39	414	9.8	1823	2 Q7KV61 DROME	Q7KV61 drosophila
40	410.5	9.7	1818	1 ENC DROME	Q6MSX1 drosophila
41	396	9.4	89	2 Q7M2N1 BOVIN	Q7M2N1 bos taurus
42	355.5	8.4	88	2 Q5FV10 RAT	Q5FV10 rattus norv
43	344.5	8.1	98	1 ARP21 MOUSE	Q9DCB4 mus musculu
44	341	8.1	558	2 Q9NPF3 CAEBL	Q9NPF3 caenorhabdi
45	318	7.5	1080	2 Q559H7 DICDI	Q559H7 dictyosteli

SQ SEQUENCE 793 AA; 86131 MW; F523AEF3E39114BE CRC64;
 Query Match 97.4%; Score 4111; DB 2; Length 793;
 Best Local Similarity 97.5%; Pred. No. 1.8e-162;
 Matches 793; Conservative 0; Mismatches 0; Indels 20; Gaps 1;

```

1 MSBQGLNDAIAEGGTGTEATPENGIVKSESLDEBEKLELORLEAQNQERRKSKSGA 60
1 MSBQGLNDAIAEGGTGTEATPENGIVKSESLDEBEKLELORLEAQNQERRKSKSGA 60
61 GKGKLTSLAVCESSARPGESLQDESILQLSSFSLSQEBDKSRKDSEREKEDKN 120
61 GKGKLTSLAVCESSARPGESLQDESILQLSSFSLSQEBDKSRKDSEREKEDKN 120
121 KDKTSKPKIRMLSKDCSQEYTDSTGIDLHEFLINTLKNNRPRMILLKMEQIIDFIAD 180
121 KDKTSKPKIRMLSKDCSQEYTDSTGIDLHEFLINTLKNNRPRMILLKMEQIIDFIAD 180
121 KDKTSKPKIRMLSKDCSQEYTDSTGIDLHEFLINTLKNNRPRMILLKMEQIIDFIAD 180
181 NNNHYKFPQMSYQOMLVHRAVAAYFGLDHNVDQTKSVIINTKSTRIPDQFCEHLKD 240
181 NNNHYKFPQMSYQOMLVHRAVAAYFGLDHNVDQTKSVIINTKSTRIPDQFCEHLKD 240
241 EKGEESQKRFILKRDNSIDKEDNQSVCQESLFEVNSRLLEDNSINCTYKKRQLEFRGN 300
241 EKGEESQKRFILKRDNSIDKEDNQSVCQESLFEVNSRLLEDNSINCTYKKRQLEFRGN 300
241 EKGEESQKRFILKRDNSIDKEDNQSVCQESLFEVNSRLLEDNSINCTYKKRQLEFRGN 300
301 RDSGSGTSGRSSSENELKMSDHQAMSTDBDSSNRMLKPMYTKTASGCTIVLTRGD 360
301 RDSGSGTSGRSSSENELKMSDHQAMSTDBDSSNRMLKPMYTKTASGCTIVLTRGD 360
281 RDSGSGTSGRSSSENELKMSDHQAMSTDBDSSNRMLKPMYTKTASGCTIVLTRGD 340
361 STSSTRTGKLSTAGSESSSSAGSSSLSTRHPLOSTPLVSGVAAGSPGVYPENGIG 420
361 STSSTRTGKLSTAGSESSSSAGSSSLSTRHPLOSTPLVSGVAAGSPGVYPENGIG 420
341 STSSTRTGKLSTAGSESSSSAGSSSLSTRHPLOSTPLVSGVAAGSPGVYPENGIG 400
421 GGVAPSTSTYILLPLEAATGIPPGSILLNPHTCQPFNPDGTALYINPPTSQOPLRSAMY 480
421 GGVAPSTSTYILLPLEAATGIPPGSILLNPHTCQPFNPDGTALYINPPTSQOPLRSAMY 480
401 GGVAPSTSTYILLPLEAATGIPPGSILLNPHTCQPFNPDGTALYINPPTSQOPLRSAMY 460
481 GGSQOQPPQOQPPQOQPPQOQPPQOQPPQOQPPQOQPPQOQPPQOQPPQOQPPQO 540
481 GGSQOQPPQOQPPQOQPPQOQPPQOQPPQOQPPQOQPPQOQPPQOQPPQOQPPQO 540
461 GGSQOQPPQOQPPQOQPPQOQPPQOQPPQOQPPQOQPPQOQPPQOQPPQOQPPQO 520
541 VSEFTQHPKRDVATQFGQWTLRSOSSGSETPEPSPGCVVYPSLSLMPQAOQPSYVIASTGQ 600
541 VSEFTQHPKRDVATQFGQWTLRSOSSGSETPEPSPGCVVYPSLSLMPQAOQPSYVIASTGQ 600
521 VSEFTQHPKRDVATQFGQWTLRSOSSGSETPEPSPGCVVYPSLSLMPQAOQPSYVIASTGQ 580
601 QLEPTGFGSGSPISQOQVLOPPSPQOQPPQOQPPQOQPPQOQPPQOQPPQOQPPQO 660
601 QLEPTGFGSGSPISQOQVLOPPSPQOQPPQOQPPQOQPPQOQPPQOQPPQOQPPQO 660
581 QLEPTGFGSGSPISQOQVLOPPSPQOQPPQOQPPQOQPPQOQPPQOQPPQOQPPQO 640
661 VQVNAARSOQMPAAQOAGYQVPLSGQOQFGGLIGVQOPOSOVILNNOGCTVQSVAMVS 720
661 VQVNAARSOQMPAAQOAGYQVPLSGQOQFGGLIGVQOPOSOVILNNOGCTVQSVAMVS 720
641 VQVNAARSOQMPAAQOAGYQVPLSGQOQFGGLIGVQOPOSOVILNNOGCTVQSVAMVS 700
721 YPTMSYQVPMTOGSGGLPQOSYQOPIMLPNQAGGSLPATGMVYCNVTPPTPQNNLRL 780
721 YPTMSYQVPMTOGSGGLPQOSYQOPIMLPNQAGGSLPATGMVYCNVTPPTPQNNLRL 780
701 YPTMSYQVPMTOGSGGLPQOSYQOPIMLPNQAGGSLPATGMVYCNVTPPTPQNNLRL 760
781 IGPHCPSSTYVPMASCRINCASMSNAGQVKF 813
761 IGPHCPSSTYVPMASCRINCASMSNAGQVKF 793
  
```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae;
 OC Homo.
 NCBI_TaxID=9606;
 RN (1)
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Testis;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Struhsberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins P.S., Wagner K.H., Schaefer C.F., Bhat N.K.,
 RA Altschul S.F., Zeeberg B., Butow K.H., Moore T., Max S.I., Wang Y., Hsieh F.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang Y., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Schetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Ketterman M., Madan A., Rodriques S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN (2)
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Testis;
 RG NIH MGCC Project;
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
 CC
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonCommercial
 CC License
 CC
 CC EMBL, BC051828; AAH51828.1; -; mRNA.
 DR SMR; Q86V31; 147-257.
 DR Ensembl; ENSG00000172995; Homo sapiens.
 DR GO; GO:0003676; F-nucleic acid binding; IEA.
 DR InterPro; IPR001374; R3H_ss_Bd.
 DR Pfam; PF01424; R3H; 1.
 DR SMART; SM00393; R3H; 1.
 DR PROSITE; PS51061; R3H; 1.
 SQ SEQUENCE 812 AA; 89196 MW; 73981A8341BB3359 CRC64;

```

1 MSBQGLNDAIAEGGTGTEATPENGIVKSESLDEBEKLELORLEAQNQERRKSKSGA 60
1 MSBQGLNDAIAEGGTGTEATPENGIVKSESLDEBEKLELORLEAQNQERRKSKSGA 60
61 GKGKLTSLAVCESSARPGESLQDESILQLSSFSLSQEBDKSRKDSEREKEDKN 120
61 GKGKLTSLAVCESSARPGESLQDESILQLSSFSLSQEBDKSRKDSEREKEDKN 120
121 KDKTSKPKIRMLSKDCSQEYTDSTGIDLHEFLINTLKNNRPRMILLKMEQIIDFIAD 180
121 KDKTSKPKIRMLSKDCSQEYTDSTGIDLHEFLINTLKNNRPRMILLKMEQIIDFIAD 180
121 KDKTSKPKIRMLSKDCSQEYTDSTGIDLHEFLINTLKNNRPRMILLKMEQIIDFIAD 180
181 NNNHYKFPQMSYQOMLVHRAVAAYFGLDHNVDQTKSVIINTKSTRIPDQFCEHLKD 240
181 NNNHYKFPQMSYQOMLVHRAVAAYFGLDHNVDQTKSVIINTKSTRIPDQFCEHLKD 240
241 EKGEESQKRFILKRDNSIDKEDNQSVCQESLFEVNSRLLEDNSINCTYKKRQLEFRGN 300
241 EKGEESQKRFILKRDNSIDKEDNQSVCQESLFEVNSRLLEDNSINCTYKKRQLEFRGN 300
241 EKGEESQKRFILKRDNSIDKEDNQSVCQESLFEVNSRLLEDNSINCTYKKRQLEFRGN 300
301 RDSGSGTSGRSSSENELKMSDHQAMSTDBDSSNRMLKPMYTKTASGCTIVLTRGD 360
301 RDSGSGTSGRSSSENELKMSDHQAMSTDBDSSNRMLKPMYTKTASGCTIVLTRGD 360
281 RDSGSGTSGRSSSENELKMSDHQAMSTDBDSSNRMLKPMYTKTASGCTIVLTRGD 340
361 STSSTRTGKLSTAGSESSSSAGSSSLSTRHPLOSTPLVSGVAAGSPGVYPENGIG 420
361 STSSTRTGKLSTAGSESSSSAGSSSLSTRHPLOSTPLVSGVAAGSPGVYPENGIG 420
341 STSSTRTGKLSTAGSESSSSAGSSSLSTRHPLOSTPLVSGVAAGSPGVYPENGIG 400
421 GGVAPSTSTYILLPLEAATGIPPGSILLNPHTCQPFNPDGTALYINPPTSQOPLRSAMY 480
421 GGVAPSTSTYILLPLEAATGIPPGSILLNPHTCQPFNPDGTALYINPPTSQOPLRSAMY 480
401 GGVAPSTSTYILLPLEAATGIPPGSILLNPHTCQPFNPDGTALYINPPTSQOPLRSAMY 460
481 GGSQOQPPQOQPPQOQPPQOQPPQOQPPQOQPPQOQPPQOQPPQOQPPQOQPPQO 540
481 GGSQOQPPQOQPPQOQPPQOQPPQOQPPQOQPPQOQPPQOQPPQOQPPQOQPPQO 540
461 GGSQOQPPQOQPPQOQPPQOQPPQOQPPQOQPPQOQPPQOQPPQOQPPQOQPPQO 520
541 VSEFTQHPKRDVATQFGQWTLRSOSSGSETPEPSPGCVVYPSLSLMPQAOQPSYVIASTGQ 600
541 VSEFTQHPKRDVATQFGQWTLRSOSSGSETPEPSPGCVVYPSLSLMPQAOQPSYVIASTGQ 600
521 VSEFTQHPKRDVATQFGQWTLRSOSSGSETPEPSPGCVVYPSLSLMPQAOQPSYVIASTGQ 580
601 QLEPTGFGSGSPISQOQVLOPPSPQOQPPQOQPPQOQPPQOQPPQOQPPQOQPPQO 660
601 QLEPTGFGSGSPISQOQVLOPPSPQOQPPQOQPPQOQPPQOQPPQOQPPQOQPPQO 660
581 QLEPTGFGSGSPISQOQVLOPPSPQOQPPQOQPPQOQPPQOQPPQOQPPQOQPPQO 640
661 VQVNAARSOQMPAAQOAGYQVPLSGQOQFGGLIGVQOPOSOVILNNOGCTVQSVAMVS 720
661 VQVNAARSOQMPAAQOAGYQVPLSGQOQFGGLIGVQOPOSOVILNNOGCTVQSVAMVS 720
641 VQVNAARSOQMPAAQOAGYQVPLSGQOQFGGLIGVQOPOSOVILNNOGCTVQSVAMVS 700
721 YPTMSYQVPMTOGSGGLPQOSYQOPIMLPNQAGGSLPATGMVYCNVTPPTPQNNLRL 780
721 YPTMSYQVPMTOGSGGLPQOSYQOPIMLPNQAGGSLPATGMVYCNVTPPTPQNNLRL 780
701 YPTMSYQVPMTOGSGGLPQOSYQOPIMLPNQAGGSLPATGMVYCNVTPPTPQNNLRL 760
781 IGPHCPSSTYVPMASCRINCASMSNAGQVKF 813
761 IGPHCPSSTYVPMASCRINCASMSNAGQVKF 793
  
```

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: July 5, 2006, 09:48:03 ; Search time 45 Seconds
(without alignments)
1738.316 Million cell updates/sec

Title: US-10-509-950-1

Perfect score: 4235
Sequence: 1 MSQGDNLNLAIEEGTEOE.....SASCRTNCASMSNAGVQVF 813

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: p1r_89:***
2: p1r1:***
3: p1r2:***
4: p1r3:***
5: p1r4:***

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3414.5	80.5	651	T42644	hypothetical prote
2	293.5	9.4	89	A34957	CAMP-regulated pro
3	282.5	6.9	2715	T13049	eyelid - fruit fly
4	279.5	6.6	838	EBWTHW	glutenin, high mol
5	274.5	6.5	848	S02262	glutenin, high mole
6	274.5	6.5	753	JC2099	glutenin, high mole
7	273.5	6.5	789	A30843	glutenin, high mole
8	267.5	6.3	791	JN0690	glutenin, high-mol
9	265.5	6.3	1006	T42731	atrophin-1 related
10	262.5	6.2	1026	T20369	hypothetical prote
11	262.5	6.2	357	S18236	omega secalin prec
12	262.5	6.2	782	T48722	hypothetical prote
13	260.5	6.2	347	T05737	probable hordein C
14	258.5	6.1	660	A24266	glutenin high mole
15	258.5	6.1	3498	T22330	hypothetical prote
16	256.5	6.1	357	S18235	omega secalin prec
17	254.5	6.0	1119	T50995	glutenin high mole
18	250.5	5.9	830	S15720	glutenin high mole
19	250.5	5.9	1184	S50832	atrophin-1 - human
20	249.5	5.9	1420	T37781	probable cytochrome
21	249.5	5.9	648	S04832	glutenin high mole
22	248.5	5.9	1184	G01763	atrophin-1 - human
23	248.5	5.9	748	T04011	hypothetical prote
24	248.5	5.9	815	JN0689	glutenin, high-mol
25	245.5	5.8	815	B30843	glutenin high mole
26	243.5	5.7	872	S33015	hypothetical prote
27	243.5	5.7	992	A31666	hypothetical prote
28	243.5	5.7	705	S18733	glutenin high mole
29	241.5	5.7	1234	T30160	hypothetical prote

30	238.5	5.6	1088	H96747	unknown protein T1
31	237.5	5.6	1307	T25563	hypothetical prote
32	235	5.5	528	I47141	gastric mucin (clo
33	235	5.5	1494	T14355	protein-tyrosine-p
34	234.5	5.5	1018	S44758	Cl4B9.6 protein -
35	234	5.5	310	T06211	C-hordein precursor
36	234	5.5	2422	T12687	ALR protein homolo
37	233.5	5.5	786	T01456	extensin homolog F
38	233	5.5	1076	T24887	hypothetical prote
39	231	5.5	802	S48529	NAB3 protein - yea
40	231	5.5	1074	T24877	hypothetical prote
41	230	5.4	620	S06733	hydroxyproline-ric
42	229	5.4	1611	T38236	hypothetical prote
43	229	5.4	2414	A54277	transcription adap
44	228.5	5.4	2129	T27431	hypothetical prote
45	226	5.3	374	T05923	glutenin low molec

ALIGNMENTS

RESULT 1
T42644
hypothetical protein DKFZP566N1047.1 - human (fragment)
C/Species: Homo sapiens (man)
C/Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C/Accession: T42644
R/Ottewald, B.; Obermayer, B.; Mewes, H.W.; Gassenhuber, J.; Wilmann, S.
submitted to the Protein Sequence Database, November 1999
A/Reference number: Z22231
A/Accession: T42644
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-651 <AAA>
A/Cross-references: UNIPROT:Q9UP93; UNIPARC:UPI000006F0B2; EMBL:AL133109
A/Experimental source: Fetal Kidney; clone DKFZP566N1047
C/genetics:
A/Note: DKFZP566N1047.1

Query Match 80.6%; Score 3414.5; DB 2; Length 651;
Best Local Similarity 99.8%; Pred. No. 1.5e-149;
Matches 651; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY	162	RDMLILKMEOEIIPIADNNHYKFPOMSSYQRMVLRVAAYFGDHNVDQTSVII	221
DB	1	RDMLILKMEOEIIPIADNNHYKFPOMSSYQRMVLRVAAYFGDHNVDQTSVII	60
QY	222	NKTSSTRIPBOHFCEHLKDEKEESQKPILRDNSSIDKEDNQVCSQESLTVENSRL	281
DB	61	NKTSSTRIPBOHFCEHLKDEKEESQKPILRDNSSIDKEDNQVCSQESLTVENSRL	119
QY	282	ESNINCETRYKKRQIFRGNRDSSGRTSGRSSSENEMLKMSDHRMWSSTDSSSRNLK	341
DB	120	ESNINCETRYKKRQIFRGNRDSSGRTSGRSSSENEMLKMSDHRMWSSTDSSSRNLK	179
QY	342	PAMTKTASFGITVLTTRGDSSTSTRSTGKLSAGSSSSSAGSSGSLSTHPTPLV	401
DB	180	PAMTKTASFGITVLTTRGDSSTSTRSTGKLSAGSSSSSAGSSGSLSTHPTPLV	239
QY	402	SGVAASPGQCVYPENGIQGVAPSTSYLLPLEAATGIPROGILNHTGQFVNPDP	461
DB	240	SGVAASPGQCVYPENGIQGVAPSTSYLLPLEAATGIPROGILNHTGQFVNPDP	299
QY	462	TPAINTPTSOQPLASAWGSGQQPPOQOBFPOQOQVQPPQMAAGPLVYOSVGLQA	521
DB	300	TPAINTPTSOQPLASAWGSGQQPPOQOBFPOQOQVQPPQMAAGPLVYOSVGLQA	359
QY	522	SSQSVQYPAVSFPPOHLLPVSPTOHFPMDVDVATOGQMTLRSOSGTEPPEPSGVYPS	581
DB	360	SSQSVQYPAVSFPPOHLLPVSPTOHFPMDVDVATOGQMTLRSOSGTEPPEPSGVYPS	419
QY	582	SLMPQAPQPSVYVASTSQQLPTGFGSGGPPISQGVLOPPSPGQFVQPPAQNPPVY	641

Db 420 SLMPDPAQPSYVIASTGQQLPTGSGSGSPISQVLPSPSGVQPPPAQMHVY 479
QY 642 YPSGGYPTSTTQGYRPMAPVQYNAORSQGMPOAAQAGYQPVLSGGQGFGLGVQOQPQ 701
Db 480 YPSGGYPTSTTQGYRPMAPVQYNAORSQGMPOAAQAGYQPVLSGGQGFGLGVQOQPQ 539
QY 702 SQNVINNOGSTPVQSVWVSYPMTSSYQVPMTQSGQLPQOSYQOPIPLPNOAGQSLPAT 761
Db 540 SQNVINNOGSTPVQSVWVSYPMTSSYQVPMTQSGQLPQOSYQOPIPLPNOAGQSLPAT 599
QY 762 GMPVYCNVTPPTPQNNRLIGPHCPSTVPVMSASCTTCASMSNMGQYKF 813
Db 600 GMPVYCNVTPPTPQNNRLIGPHCPSTVPVMSASCTTCASMSNMGQYKF 651

RESULT 2
A34957
CAMP-regulated phosphoprotein 21 - bovine
N:Alternate names: ARPP-21
C:Species: Bos primigenius taurus (cattle)
C>Date: 13-Jul-1990 #sequence_revision 13-Jul-1990 #text_change 09-Jul-2004
C:Accession: A34957, A34956
R:Kurthare, T.; Ehrlich, M.E.; Horiuchi, J.; Nasu, T.; Greengard, P.
J. Neurosci. 9, 3638-3644, 1989
A:Title: ARPP-21, a cyclic AMP-regulated phosphoprotein enriched in dopamine-innervated
A:Reference number: A34957, NCID:90011251, PMID:2552037
A:Accession: A34957
A>Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-89 <KUR>
A:Cross-references: UNIPROT:Q7M2N1; UNIPARC:UP1000017C4B9
J. Williams, K.R.; Hemmings Jr., H.C.; LoPresti, M.B.; Greengard, P.
J. Neurosci. 9, 3631-3637, 1989
A:Title: ARPP-21, a cyclic AMP-regulated phosphoprotein enriched in dopamine-innervated
A:Reference number: A34956, NCID:90011250, PMID:2552036
A:Accession: A34956
A:Molecule type: protein
A:Residues: 2-25, 'D', 27-89 <WIL>
A:Cross-references: UNIPARC:UP1000017C4BA
A>Note: this sequence was obtained from isoform ARPP-21B
C:Keyword: acetylated amino end; phosphoprotein 21 #status predicted <WAT>
F1/2-89/Product: CAMP-regulated phosphoprotein 21 #status experimental
F1/2/Modified site: acetylated amino end (Ser) (in mature form) #status experimental
F1/56/Binding site: phosphate (Ser) (covalent) (by CAMP-dependent kinase) #status experiment

Query Match 9.4%; Score 396; DB 2; Length 89;
Best Local Similarity 90.8%; Pred. No. 3.4e-12; Indels 0; Gaps 0;
Matches 79; Conservative 3; Mismatches 5;

QY 1 MSEGQNLQAIAEEGTEQETATPENGIVKESLDEBEKLELQRLAQNQERRKSKGA 60
Db 1 MSEGQNLQAIAEEGTEQETATPENGIVKESLDEBEKLELQRLAQNQERRKSKGA 60
QY 61 GKGLTRSLAVCESSARPGESLQDQ 87
Db 61 GKGLTRSLAVCESSARPGESLQDQ 87

RESULT 3
T13049
eyelid - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C>Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C:Accession: T13049
R:Firestein, J.E.; Luk, A.; Rubin, G.M.; Heberlein, U.
Submitted to the EMBL Data Library, March 1998
A:Reference number: Z17592
A:Accession: T13049
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-2715 <TRB>
A:Cross-references: UNIPROT:Q9IN94; UNIPARC:UP1000007D87E; EMBL:AF053091; NCID:92981220;
C:Genetics:

A:Gene: eid
A:Cross-references: FlyBase:FBgn0003013
C:Function:
A:Description: could act as a transcription factor antagonistic to the Wg pathway
C:Keywords: DNA binding

Query Match 6.9%; Score 293.5; DB 2; Length 2715;
Best Local Similarity 22.0%; Pred. No. 8.7e-06;
Matches 215; Conservative 71; Mismatches 334; Indels 357; Gaps 45;

QY 4 QEDLNQALIEEGTQ--ETATPENGIVKESLDEBEKLELQRLAQNQERRKSKGA 60
Db 835 QGGQPPQGGASGGPESGPEHISODNGI-----SSS 865

QY 61 GKGLTRSLAVCESSARPGESLQDQESIHQLSFSLSQED-----KSRKDSERE 114
Db 866 GPTGAAGMAHVVTSVTTGPDGISM-DEVQOSTLSNASHAGBDDPQCTTPKSKRNDP--Y 922

QY 115 KEKDKNKDKTSEKPKIRMLSKDCSQRYTSTGIDLHEFLINTLKNNSRDMILKMEQEI 174
Db 923 SQSHLAPSTSDHPVVMHGGGGEYDMSSP----- 954

QY 175 IDPLADNNHKKFPQMSYQRMLYHRVAAYFGLDHNDQTKSVIINKTSSTRIPQRF 234
Db 955 -----FNWPRPAGSPQVFNR-----VVPQEPF 977

QY 235 CEHLDEKGEESQKPILRDNSIDKEDNOSVQSESLFVENSRLLEDNINCNETKY 292
Db 978 RSTITTTKSDSLCKLYEMDDN-----PDRGMLDKLRPMERRPITACPTISKQPLD 1032

QY 293 KQOLFKNRDSGSGRTSGRSSSENLKMSDQRAWSTSDSSNNRLKPMATKTASFGG 352
Db 1033 LYRLIYVERGGFVETKSKT-----WKDIAGLIGIGASSAAVTLRKHYK----- 1080

QY 353 ITVLT-----RGD-----STSTRSTGKLSKASB-----SSSAGSGSL 388
Db 1081 -NLTTFECHFDGDDIDPLPIQVENVGSKKTKAKASVSPSGGHLDACTNSTGSSNSQ 1139

QY 389 SRTHPLOGSTP--LVSGVAAGSPGCVPY-----BENGI 419
Db 1140 DSEFPAPGAPMAALDGY-PGYFGSPYFGASGPPQDYATAGQMQRPPSNNQTHRPGA 1198

QY 420 GGQVAPSSSTYLLPLE--AATGIPPSILNPHNQ--PVPNDGTPAI-----Y 466
Db 1199 AAAVAAGNISVSNPEEDPIAAGGPGS-GTGGPGQGGPRAAGSAGAVGAVGGGPPQ 1257

QY 467 NPPTSQQLRSM--VGSGQQPPOQO--PSPPQOQVQ-----PPQ 505
Db 1258 HPPPHSPHTAAQQAAGHQHQHPOHQHGLPQPPPPQOQGGQGGQPPSVGGGPPAP 1317

QY 506 QWAGPLVTVSGQLQASSQSVQYPAVS--FPP--QHLLPVS--PTQHFMRDQVATQFG 558
Db 1318 QQHGP-----GQVPPSPQAHVPAAGARYPPGSGSYPTPVSKTTPSSPISQAGAGQYQ 1371

QY 559 QWTLNR-----QSSGETPE-----PPSGP-----VYPSLSMPQ-- 587
Db 1372 SSDQVYATGPPPGQPFQGGPQGYPPQNRNMVPPYGPGEGBAPPTGANGYGFGRSPYQPP 1431

QY 588 --AQPSYIATGQQLPTGFGSGSGP-----PISQVLOPPSPQGFVQPPPAQM 637
Db 1432 GGPQPPQTQVAGGP--PAGGAPGAPSSGAYPTGRSQDDYQYPPD-----OSPPRRH 1483

QY 638 PLYVYVSGGYF-----TSTTQYRPM----- 658
Db 1484 PDEIKDSQYPPGYNARPOIYGAMQSGTQGYRQYPSPPAPQWNGAPPPGAAPPPAPHG 1543

QY 659 APVQYNA-----QSSQMPQAQAGYQPVUSGQGFGLGVQO 698
Db 1544 PPQQPAGVAVQWDHRRYPPQGGPPPPQOQOQPPQOQPPYQOV-AGPG-----QQ 1595

QY 699 PPOS-QNVINNOGGTGVQSVWVSYPMTSSYQVPMTQSGO-----GLP-----QOSYQO--PI 747
Db 1596 PPOAPPQMAQNNPQGTAGSGIA--PPGSLRPPSGGQGNRMGMVAQOQOQSGQGGVQ 1653

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: July 5, 2006, 09:49:57 ; Search time 99 Seconds

(Without alignments)
3803.977 Million cell updates/sec

Title: US-10-509-950-1

Perfect score: 4235
Sequence: 1 MSQGDNLQAIABEGGTGE.....SASCRITVCAASNAGWVKF 813

Scoring table: BIOSOM62

Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database:

Published Applications AA Main:
1: /BMC_Celerra_SIDS3/Prodata/2/pubppaa/US07_PUBCOMB.pep.*
2: /BMC_Celerra_SIDS3/Prodata/2/pubppaa/US08_PUBCOMB.pep.*
3: /BMC_Celerra_SIDS3/Prodata/2/pubppaa/US09_PUBCOMB.pep.*
4: /BMC_Celerra_SIDS3/Prodata/2/pubppaa/US10A_PUBCOMB.pep.*
5: /BMC_Celerra_SIDS3/Prodata/2/pubppaa/US10B_PUBCOMB.pep.*
6: /BMC_Celerra_SIDS3/Prodata/2/pubppaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4235	100.0	813	4 US-10-112-372-6	Sequence 6, App1
2	4235	100.0	813	5 US-10-509-950-1	Sequence 1, App1
3	4208	99.4	847	4 US-10-112-372-2	Sequence 2, App1
4	4111	97.1	793	4 US-10-112-372-8	Sequence 8, App1
5	4084	96.4	827	4 US-10-112-372-4	Sequence 4, App1
6	3421	80.8	807	4 US-10-112-372-11	Sequence 11, App1
7	3414.5	80.6	651	4 US-10-112-372-13	Sequence 13, App1
8	3398.5	80.2	648	4 US-10-112-372-12	Sequence 12, App1
9	2513.5	59.4	635	5 US-10-450-763-35346	Sequence 35346, A
10	2047	48.3	438	5 US-10-450-763-32449	Sequence 32449, A
11	1831.5	43.2	1099	5 US-10-756-149-5469	Sequence 5469, A
12	1493	35.3	976	4 US-10-144-194A-92	Sequence 92, App1
13	1493	35.3	976	5 US-10-491-566-92	Sequence 92, App1
14	1455	34.3	1027	5 US-10-490-318-10	Sequence 10, App1
15	1451	34.3	958	4 US-10-144-194A-90	Sequence 10, App1
16	1451	34.3	958	5 US-10-491-566-90	Sequence 90, App1
17	1420.5	34.1	433	5 US-10-450-763-32447	Sequence 32447, A
18	1020	24.1	366	5 US-10-450-763-35345	Sequence 35345, A
19	944	22.3	371	5 US-10-450-763-32448	Sequence 32448, A
20	473	11.2	107	4 US-10-104-047-3736	Sequence 3736, App
21	473	11.2	107	6 US-11-072-512-3736	Sequence 3736, App
22	430	10.2	89	4 US-10-112-372-10	Sequence 10, App1
23	430	10.2	89	4 US-10-112-372-14	Sequence 14, App1
24	410.5	9.7	1805	6 US-11-097-143-22578	Sequence 22578, A
25	346.5	8.2	94	5 US-10-450-763-32444	Sequence 32444, A
26	346.5	8.2	94	5 US-10-450-763-35343	Sequence 35343, A
27	344.5	8.1	88	4 US-10-112-372-15	Sequence 15, App1

28	311.5	7.4	78	3 US-09-864-408A-5512	Sequence 5512, App
29	308	7.3	79	5 US-10-450-763-32445	Sequence 32445, A
30	308	7.3	79	5 US-10-450-763-35344	Sequence 35344, A
31	293	6.9	2004	5 US-10-469-469-250	Sequence 250, App
32	291	6.9	2703	6 US-11-097-143-16689	Sequence 16689, A
33	286	6.8	1390	4 US-10-092-900A-224	Sequence 224, App
34	274.5	6.5	1566	5 US-10-491-545A-63	Sequence 63, App1
35	268	6.3	3394	4 US-10-408-765A-529	Sequence 529, App
36	265	6.3	1026	5 US-10-415-656-2	Sequence 2, App1
37	263.5	6.2	926	6 US-11-097-143-22197	Sequence 22197, A
38	260	6.1	1250	4 US-10-112-372-249	Sequence 249, App
39	259.5	6.1	1240	4 US-10-369-493-4031	Sequence 4031, App
40	256.5	6.1	2759	5 US-10-450-763-47507	Sequence 47507, A
41	255	6.0	1189	4 US-10-408-765A-2272	Sequence 2272, App
42	251.5	5.9	2285	5 US-10-773-446-101	Sequence 101, App
43	251	5.9	2280	6 US-11-097-143-11742	Sequence 11742, A
44	249	5.9	3080	6 US-11-097-143-21423	Sequence 21423, A
45	248	5.9	1790	6 US-11-097-143-42009	Sequence 42009, A

ALIGNMENTS

RESULT 1
US-10-112-372-6
; Sequence 6, Application US/10112372
; Publication No. US20030186249A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies, Inc.
; TITLE OR INVENTION: Human TRAP Genes and Polypeptides
; FILE REFERENCE: 16U 105 R1
; CURRENT APPLICATION NUMBER: US/10/112,372
; CURRENT FILING DATE: 2002-04-01
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 813
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-112-372-6

Query Match 100.0%; Score 4235; DB 4; Length 813;
Best Local Similarity 100.0%; Pred. No. 8.2e-234;
Matches 813; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MSQGDNLQAIABEGGTGEQTATPENGIVKSESLDEEKEKLQRLAQNQERKKS	60
DB	1	MSQGDNLQAIABEGGTGEQTATPENGIVKSESLDEEKEKLQRLAQNQERKKS	60
QY	61	GKGLTRSLAVCESSAPGSESLDQESIHQLSSFSISOEDKSKKDSEREKEDKN	120
DB	61	GKGLTRSLAVCESSAPGSESLDQESIHQLSSFSISOEDKSKKDSEREKEDKN	120
QY	121	KKTESEKPIKRLMSKDCSQEYTDSTGIDHAEFLINTLKNNSRDRMLTKMEOBII	180
DB	121	KKTESEKPIKRLMSKDCSQEYTDSTGIDHAEFLINTLKNNSRDRMLTKMEOBII	180
QY	181	NNNHKKKPPQMSYQRMVHRVAAYFGDLNHDQCKSYIINTKSTRIPKPCFCHLKD	240
DB	181	NNNHKKKPPQMSYQRMVHRVAAYFGDLNHDQCKSYIINTKSTRIPKPCFCHLKD	240
QY	241	EKGESQKRFILKRNKSGIDKEDNOSVCSQESLFEYNSRLLEDNSNENETKPKRQ	300
DB	241	EKGESQKRFILKRNKSGIDKEDNOSVCSQESLFEYNSRLLEDNSNENETKPKRQ	300
QY	301	RDGSGRTGSGRSSSENEIKMWDHQRAWSSTDSNNRLKAPMTKTASFGGITVLT	360
DB	301	RDGSGRTGSGRSSSENEIKMWDHQRAWSSTDSNNRLKAPMTKTASFGGITVLT	360
QY	361	STSSRSRTGSKSKASSESSSSAGSSGSLSRTHPLQSTGLVGVAGSGCVYPENGG	420
DB	361	STSSRSRTGSKSKASSESSSSAGSSGSLSRTHPLQSTGLVGVAGSGCVYPENGG	420

```
QY 421 GQVAPSTSYILLPLEAATGIPGSIILNPHTCQPFVNPDPGTALVNPPTSQOPLASAMV 480
DB 421 GQVAPSTSYILLPLEAATGIPGSIILNPHTCQPFVNPDPGTALVNPPTSQOPLASAMV 480
QY 481 GGSQQQPPOQPSPPQOQVQPPQPMAGPLVTQSVQGLQASSQSQVQPAVSFPPOHLLP 540
DB 481 GGSQQQPPOQPSPPQOQVQPPQPMAGPLVTQSVQGLQASSQSQVQPAVSFPPOHLLP 540
QY 541 VSTPHFPMKDDVAATQFGQMTLSRSSGGETPEPPSGPVYSSILMPQPAQOQPSVIASTGQ 600
DB 541 VSTPHFPMKDDVAATQFGQMTLSRSSGGETPEPPSGPVYSSILMPQPAQOQPSVIASTGQ 600
QY 601 QLEPTGFSGSGPPIISQOVLQPPSPQGFVQOQPPAPMPVYVYPSGQYPTSTTQOYRPMAP 660
DB 601 QLEPTGFSGSGPPIISQOVLQPPSPQGFVQOQPPAPMPVYVYPSGQYPTSTTQOYRPMAP 660
QY 661 VOYNAGRSQQMPQAAQOAGYQPVLSGQGFQGLIGVQOQPSQNVINNQGTFVQSVYMS 720
DB 661 VOYNAGRSQQMPQAAQOAGYQPVLSGQGFQGLIGVQOQPSQNVINNQGTFVQSVYMS 720
QY 721 YPTMSSYQVPMTOGSGGLPQOQSVQOPIMLPNOAGGSLPATGMPVYCNVTPPTPONNRL 780
DB 721 YPTMSSYQVPMTOGSGGLPQOQSVQOPIMLPNOAGGSLPATGMPVYCNVTPPTPONNRL 780
QY 781 IGPCHCSSTVPVMSASCRINCASMSNAGQVXF 813
DB 781 IGPCHCSSTVPVMSASCRINCASMSNAGQVXF 813
```

```
RESULT 2
US-10-509-950-1
; Sequence 1, Application US/10509950
; Publication No. US20060024305A1
; GENERAL INFORMATION:
; APPLICANT: Evotec Neurosciences GmbH
; TITLE OF INVENTION: CAMP-Regulated Phosphoprotein for Diagnostic and
; FILE REFERENCE: 020880ep
; CURRENT APPLICATION NUMBER: US/10/509,950
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 813
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-509-950-1
```

```
Query Match 100.0%; Score 4235; DB 5; Length 813;
Best Local Similarity 100.0%; Pred. No. 8.2e-234;
Matches 813; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MSEQGLNOLAIEBGGTEOETATPENGIVKSESILDEBEKLELORLEAQNQERRKSKSGA 60
DB 1 MSEQGLNOLAIEBGGTEOETATPENGIVKSESILDEBEKLELORLEAQNQERRKSKSGA 60
QY 61 GKGLTRSLAVCESSARPGESLQOQESIHQLSSFSLSQEBDKSRKDSEREKEKDXN 120
DB 61 GKGLTRSLAVCESSARPGESLQOQESIHQLSSFSLSQEBDKSRKDSEREKEKDXN 120
QY 121 KQKTSSEKPKIRMLSKDCSOEYTDSTGIDLHEFLINTLKNSRDRMILLKXEOETIDFIAD 180
DB 121 KQKTSSEKPKIRMLSKDCSOEYTDSTGIDLHEFLINTLKNSRDRMILLKXEOETIDFIAD 180
QY 181 NNNHYKKFPOMSSYQRMVLRVAAYFGLDHNVDQTKSVIINTKSTRIFEOQFCEHLKD 240
DB 181 NNNHYKKFPOMSSYQRMVLRVAAYFGLDHNVDQTKSVIINTKSTRIFEOQFCEHLKD 240
QY 241 EKSEESQKRFILKRDNSIDKEDNOSVCSQESLFEVNSRLLEDSDNICNETYKKGQULFRGN 300
DB 241 EKSEESQKRFILKRDNSIDKEDNOSVCSQESLFEVNSRLLEDSDNICNETYKKGQULFRGN 300
QY 301 RDGSGRTSGROSSENELKMSDHQRAMSSTDSDSNRLKPMATKTASFGGTLVLTTRGD 360
```

```
DB 301 RDGSGRTSGROSSENELKMSDHQRAMSSTDSDSNRLKPMATKTASFGGTLVLTTRGD 360
QY 361 STSSTSTGKLSKAGSESSSSSAGSSLSRTHPPLQSTPLVSGVAAGSPCCVYPPENGIG 420
DB 361 STSSTSTGKLSKAGSESSSSSAGSSLSRTHPPLQSTPLVSGVAAGSPCCVYPPENGIG 420
QY 421 GQVAPSTSYILLPLEAATGIPGSIILNPHTCQPFVNPDPGTALVNPPTSQOPLASAMV 480
DB 421 GQVAPSTSYILLPLEAATGIPGSIILNPHTCQPFVNPDPGTALVNPPTSQOPLASAMV 480
QY 481 GGSQQQPPOQPSPPQOQVQPPQPMAGPLVTQSVQGLQASSQSQVQPAVSFPPOHLLP 540
DB 481 GGSQQQPPOQPSPPQOQVQPPQPMAGPLVTQSVQGLQASSQSQVQPAVSFPPOHLLP 540
QY 541 VSTPHFPMKDDVAATQFGQMTLSRSSGGETPEPPSGPVYSSILMPQPAQOQPSVIASTGQ 600
DB 541 VSTPHFPMKDDVAATQFGQMTLSRSSGGETPEPPSGPVYSSILMPQPAQOQPSVIASTGQ 600
QY 601 QLEPTGFSGSGPPIISQOVLQPPSPQGFVQOQPPAPMPVYVYPSGQYPTSTTQOYRPMAP 660
DB 601 QLEPTGFSGSGPPIISQOVLQPPSPQGFVQOQPPAPMPVYVYPSGQYPTSTTQOYRPMAP 660
QY 661 VOYNAGRSQQMPQAAQOAGYQPVLSGQGFQGLIGVQOQPSQNVINNQGTFVQSVYMS 720
DB 661 VOYNAGRSQQMPQAAQOAGYQPVLSGQGFQGLIGVQOQPSQNVINNQGTFVQSVYMS 720
QY 721 YPTMSSYQVPMTOGSGGLPQOQSVQOPIMLPNOAGGSLPATGMPVYCNVTPPTPONNRL 780
DB 721 YPTMSSYQVPMTOGSGGLPQOQSVQOPIMLPNOAGGSLPATGMPVYCNVTPPTPONNRL 780
QY 781 IGPCHCSSTVPVMSASCRINCASMSNAGQVXF 813
DB 781 IGPCHCSSTVPVMSASCRINCASMSNAGQVXF 813
```

```
RESULT 3
US-10-112-372-2
; Sequence 2, Application US/10112372
; Publication No. US20030186249A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies, Inc.
; TITLE OF INVENTION: Human Tarpp Genes and Polypeptides
; FILE REFERENCE: 16U 105 R1
; CURRENT APPLICATION NUMBER: US/10/112,372
; CURRENT FILING DATE: 2002-04-01
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 847
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-112-372-2
```

```
Query Match 99.4%; Score 4208; DB 4; Length 847;
Best Local Similarity 96.0%; Pred. No. 3e-232;
Matches 813; Conservative 0; Mismatches 0; Indels 34; Gaps 1;
```

```
QY 1 MSEQGLNOLAIEBGGTEOETATPENGIVKSESILDEBEKLELORLEAQNQERRKSKSGA 60
DB 1 MSEQGLNOLAIEBGGTEOETATPENGIVKSESILDEBEKLELORLEAQNQERRKSKSGA 60
QY 61 GKGLTRSLAVCESSARPGESLQOQESIHQLSSFSLSQEBDKSRKDSEREKEKDXN 120
DB 61 GKGLTRSLAVCESSARPGESLQOQESIHQLSSFSLSQEBDKSRKDSEREKEKDXN 120
QY 121 KQKTSSEKPKIRMLSKDCSOEYTDSTGIDLHEFLINTLKNSRDRMILLKXEOETIDFIAD 180
DB 121 KQKTSSEKPKIRMLSKDCSOEYTDSTGIDLHEFLINTLKNSRDRMILLKXEOETIDFIAD 180
QY 181 NNNHYKKFPOMSSYQRMVLRVAAYFGLDHNVDQTKSVIINTKSTRIFEOQFCEHLKD 240
DB 181 NNNHYKKFPOMSSYQRMVLRVAAYFGLDHNVDQTKSVIINTKSTRIFEOQFCEHLKD 240
```


Db 241 GGTAAAGTCTTCAGATCAAGATCAATTCATTACAGCTTTCAGTTTTCAGCTG 300
Qy 301 CAAGAGAGATTAATCTAGAGAAAGTACTCTGAAGAGAGAGAGAGAGAGATTAAC 360
Db 301 CAAGAGAGATTAATCTAGAGAAAGTACTCTGAAGAGAGAGAGAGAGATTAAC 360
Qy 361 AAAGATTAACCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Db 361 AAAGATTAACCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Qy 421 TACAGAGATCTACAGAGATCAAGATCAAGATCTGATTAACATTAAGAGATTAAT 480
Db 421 TACAGAGATCTACAGAGATCAAGATCAAGATCTGATTAACATTAAGAGATTAAT 480
Qy 481 TCCAGAGAGAGAGATCAAGATCAAGATCAAGAGAGAGAGAGAGAGAGAGAGAG 540
Db 481 TCCAGAGAGAGAGATCAAGATCAAGATCAAGAGAGAGAGAGAGAGAGAGAGAG 540
Qy 541 AACATTAATCAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Db 541 AACATTAATCAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Qy 601 CGAGTGCAGCTTATTTTGAATGATCAATGATGATCAAGAGAGAGAGAGAGAGATC 660
Db 601 CGAGTGCAGCTTATTTTGAATGATCAATGATGATCAAGAGAGAGAGAGAGATC 660
Qy 661 ATCAAGAT 720
Db 661 ATCAAGAT 720
Qy 721 GAAAAAGTGAAGAT 780
Db 721 GAAAAAGTGAAGAT 780
Qy 781 AAAGAT 840
Db 781 AAAGAT 840
Qy 841 TTGAT 900
Db 841 TTGAT 900
Qy 901 AGAGATGCTCAGAT 960
Db 901 AGAGATGCTCAGAT 960
Qy 961 TGGTCTGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1020
Db 961 TGGTCTGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1020
Qy 1021 AAGCCGAT 1080
Db 1021 AAGCCGAT 1080
Qy 1081 AGCATTCTCAGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1140
Db 1081 AGCATTCTCAGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1140
Qy 1141 AGTCAAGAT 1200
Db 1141 AGTCAAGAT 1200
Qy 1201 GTCTCAGAT 1260
Db 1201 GTCTCAGAT 1260
Qy 1261 GGGCAGAT 1320
Db 1261 GGGCAGAT 1320
Qy 1321 ATCCGAT 1380
Db 1321 ATCCGAT 1380

Db 1381 ATCCGAT 1380
Qy 1381 GGAATCTCTGAT 1440
Db 1381 GGAATCTCTGAT 1440
Qy 1441 GGGCAGAT 1500
Db 1441 GGGCAGAT 1500
Qy 1501 CAGCAGAT 1560
Db 1501 CAGCAGAT 1560
Qy 1561 GCTTCTCTCAGAT 1620
Db 1561 GCTTCTCTCAGAT 1620
Qy 1621 GTGCTCAGAT 1680
Db 1621 GTGCTCAGAT 1680
Qy 1681 ACCCTGAT 1740
Db 1681 ACCCTGAT 1740
Qy 1741 TCTCTCTCTGAT 1800
Db 1741 TCTCTCTCTGAT 1800
Qy 1801 CAGCTTCTCAGAT 1860
Db 1801 CAGCTTCTCAGAT 1860
Qy 1861 CCCCCTCTCAGAT 1920
Db 1861 CCCCCTCTCAGAT 1920
Qy 1921 TATTAAGAT 1980
Db 1921 TATTAAGAT 1980
Qy 1981 GTTCAAGAT 2040
Db 1981 GTTCAAGAT 2040
Qy 2041 CAGCAGAT 2100
Db 2041 CAGCAGAT 2100
Qy 2101 CAGAT 2160
Db 2101 CAGAT 2160
Qy 2161 TACCAAGAT 2220
Db 2161 TACCAAGAT 2220
Qy 2221 CAGTCAAGAT 2280
Db 2221 CAGTCAAGAT 2280
Qy 2281 ACTGAT 2340
Db 2281 ACTGAT 2340
Qy 2341 ATTGAT 2400
Db 2341 ATTGAT 2400
Qy 2401 TGTGAT 2442
Db 2401 TGTGAT 2442